

Human albumin (HA)
Human albumin (Hh)
Human albumin (Hb)
Human serum albumin
Human serum albumin
B lymphocyte stimula
Human B lymphocyte
Mature form of hum
Recombinant human
Recombinant human
pre human serum al
Chimeric human ser
Recombinant human
Recombinant human
Prepro-HSA-G-CSF c
G-CSF-(GL1)-4-HSA
HSA-VWF(470-713) f
Sequence of mature
Mature protein of
Human serum albumin
Sequence of human
Human serum albumin
Cancer metastasis
Human serum albumin
Human serum albumin
Human albumin. Ho
Myosin light chain
Human serum albumin
Human serum albumin
Human serum albumin
Human serum albumin
Human serum albumin
Sequence of prepro
Human serum albumin
HSA. Pichia pastor
HSA-gamma RII f
Human serum albumin

PS Disclosure; fig 2; 20pp. English.

CC Mature protein of human serum albumin (see corresp. AAN90128).
CC Used to make new N-terminal fragments which are used as plasma
CC expanders, or as substitutes for HSA or BSA, in tissue culture
CC media.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
DB 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
QY 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
QY 121 DVMTAFHNEETFLKLYEIAARRHPYFVAPPELLFEAKRYKAAFTTECCQAADRAACLLP 180
DB 121 DVMTAFHNEETFLKLYEIAARRHPYFVAPPELLFEAKRYKAAFTTECCQAADRAACLLP 180
QY 181 KLDELROGKASSAKORLKASLOKFGERAFAKMAVAARLSORFPAEFAEYSKLVTDLTK 240
DB 181 KLDELROGKASSAKORLKASLOKFGERAFAKMAVAARLSORFPAEFAEYSKLVTDLTK 240
QY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
QY 301 DLPSIADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSIADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLTKC 360
QY 361 CAADPHCYAKVDEFPVLEEPPONLIKONCELEFOLGEYKFFONALLVRYTKKVPVST 420
DB 361 CAADPHCYAKVDEFPVLEEPPONLIKONCELEFOLGEYKFFONALLVRYTKKVPVST 420
QY 421 PTLVEVSNTLGKVGSKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
DB 421 PTLVEVSNTLGKVGSKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
QY 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKERQIKQIALVELVYKHKPKAT 540
DB 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKERQIKQIALVELVYKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEAGKTLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEAGKTLVAASQAALGL 585

RESULT 2
AAR05318
ID AAR05318 standard; protein; 585 AA.
AC AAR05318;
XX 08-OCT-1990 (first entry)
XX Human serum albumin gene product.
XX Human serum albumin; HSA-A; yeast; ds.
XX Homo sapiens.
XX JP02117384-A.
XX 01-MAY-1990.
XX 26-OCT-1988; 88JP-0268302.

XX 26-OCT-1988; 88JP-0268302.
XX (TOFU) TOA NENRYO KOGYO KK.
XX WPI: 1990-176228/23.
XX N-PSDB: AAO04719.

PT Human serum albumin prepn. by yeast host -
PT by culturing transformed plasmid yeast to produce serum, and
PT removing it.

PS Disclosure; ; pp; Japanese.

CC Mature HSA-A may be produced using the sequence incorporated into a
CC plasmid vector with suitable controllers, and transferred to a yeast
CC expression system.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
DB 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
QY 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
QY 121 DVMTAFHNEETFLKLYEIAARRHPYFVAPPELLFEAKRYKAAFTTECCQAADRAACLLP 180
DB 121 DVMTAFHNEETFLKLYEIAARRHPYFVAPPELLFEAKRYKAAFTTECCQAADRAACLLP 180
QY 181 KLDELROGKASSAKORLKASLOKFGERAFAKMAVAARLSORFPAEFAEYSKLVTDLTK 240
DB 181 KLDELROGKASSAKORLKASLOKFGERAFAKMAVAARLSORFPAEFAEYSKLVTDLTK 240
QY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
QY 301 DLPSIADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSIADFVESKDVCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLTKC 360
QY 361 CAADPHCYAKVDEFPVLEEPPONLIKONCELEFOLGEYKFFONALLVRYTKKVPVST 420
DB 361 CAADPHCYAKVDEFPVLEEPPONLIKONCELEFOLGEYKFFONALLVRYTKKVPVST 420
QY 421 PTLVEVSNTLGKVGSKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
DB 421 PTLVEVSNTLGKVGSKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
QY 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKERQIKQIALVELVYKHKPKAT 540
DB 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKERQIKQIALVELVYKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEAGKTLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEAGKTLVAASQAALGL 585

RESULT 3
AAR08457
ID AAR08457 standard; Protein; 585 AA.
AC AAR08457;
XX 16-APR-1991 (first entry)

DE Human serum albumin.
 XX HSA; folding; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Region 123..303
 FT /label= A
 FT Region 1..303
 FT /label= B
 FT Region 123..585
 FT /label= C
 XX
 XX JP02227079-A.
 PN
 PD 25-AUG-1989.
 XX
 XX 10-SEP-1990: 90JP-0250926.
 PF
 XX 06-OCT-1988: 88JP-0250926.
 PR
 XX (TOFU) TONEN CORP.
 PA
 XX WPI: 1990-317325/42.
 DR N-PSDB: AAQ06099.
 XX
 XX
 PT New human serum albumin fragments - used to bond medicines and for
 PT stable folding of protein(s).
 PS
 XX Claim 1: Fig 8: 24pp; Japanese.
 XX
 CC Fragments A-C of HSA are expressed as fusion proteins with the
 CC signal peptide of E. coli alkaline phosphatase. The fragments are
 CC selected for their specific properties. The C-terminal truncated
 CC fragment, B, does not bind long-chain fatty acids but does bind to
 CC various medicines at the central region. The N-terminal truncated
 CC fragment, C, has good stability in protein folding. The central
 CC segment, A, has characteristics of both B and C.
 CC See also AAQ06096-006098.
 CC
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 11: Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
 DB 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPBERNECFQHKDNDNPLRLVPRVY 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPBERNECFQHKDNDNPLRLVPRVY 120
 QY 121 DWACSTAFHNEETFLKLYETARRRHPYAPAPLLEFARKYKAFTECCOAAKACCLIP 180
 DB 121 DWACSTAFHNEETFLKLYETARRRHPYAPAPLLEFARKYKAFTECCOAAKACCLIP 180
 QY 181 KIDELRDEKASSAKQRLKASLQKGERAFKAMAVARLSQRPKAEFAEVSKLVTDLTK 240
 DB 181 KIDELRDEKASSAKQRLKASLQKGERAFKAMAVARLSQRPKAEFAEVSKLVTDLTK 240
 QY 241 VITECHGDLLECADRADLAKYICENODSISSKLECECEKPLLEKSHCIAEYENDMPA 300
 DB 241 VITECHGDLLECADRADLAKYICENODSISSKLECECEKPLLEKSHCIAEYENDMPA 300
 QY 301 DLPSLAADPVESKDVCKNTAAEKDVFLEGMFLYEYARRHPDYSVLLRLAKYETLLEK 360
 DB 301 DLPSLAADPVESKDVCKNTAAEKDVFLEGMFLYEYARRHPDYSVLLRLAKYETLLEK 360
 QY 361 CAADHHECYAKVDEFFKPLVEEPQNLINQNCLEFQDLGEYKFNOMALVRYTKVQVST 420
 DB 361 CAADHHECYAKVDEFFKPLVEEPQNLINQNCLEFQDLGEYKFNOMALVRYTKVQVST 420

DB 361 CAADHHECYAKVDEFFKPLVEEPQNLINQNCLEFQDLGEYKFNOMALVRYTKVQVST 420
 QY 421 PTLVEVSRNLGKSGSCCKHPEAKRMPCAEDYLSVVLNQLCVLHETTPVSDRTKCTES 480
 DB 421 PTLVEVSRNLGKSGSCCKHPEAKRMPCAEDYLSVVLNQLCVLHETTPVSDRTKCTES 480
 QY 481 LVNRRPCFSALVEDETVPEKEFNAETFTFHADICTSEKERQTKOTALVELVKKRPAT 540
 DB 481 LVNRRPCFSALVEDETVPEKEFNAETFTFHADICTSEKERQTKOTALVELVKKRPAT 540
 QY 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 585
 DB 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 585
 RESULT 4
 AAR80301
 ID AAR80301 standard; Protein; 585 AA.
 XX
 AC AAR80301;
 XX
 DT 17-JAN-1996 (first entry)
 XX
 DE Human serum albumin.
 XX
 KM Serum albumin; HSA; aspartyl protease-3; Yap3p;
 KM Saccharomyces cerevisiae.
 XX
 OS Homo sapiens.
 XX
 XX W09523857-A1.
 XX
 PD 08-SEP-1995.
 XX
 PF 01-MAR-1995: 95WO-GB00434.
 XX
 PR 05-MAR-1994: 94GB-0004270.
 XX
 PA (DELTA) DELTA BIOTECHNOLOGY LTD.
 XX
 XX Gilbert SC, Kerry-Williams SM;
 PI
 DR WPI: 1995-320572/41.
 DR N-PSDB: AAQ98695.
 XX
 PT Yeast with reduced levels of aspartyl protease 3 proteolytic
 PT activity - used to secrete human albumin without prodn. of the 45
 PT kd fragment
 XX
 PS Example 1; Page 26-28; 50pp; English.
 XX
 CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of
 CC endoproteases in the generation of a 45 kDa albumin fragment obtd.
 CC when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;
 CC I407A, I408Y, V409A, and R410A, K413Q, K414Q. The latter set of
 CC mutations, especially, improved stability of HSA to yeast Yap3p
 CC proteolytic cleavage, allowing increased prodn. of recombinant HSA.
 CC
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 16: Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
 DB 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPBERNECFQHKDNDNPLRLVPRVY 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPBERNECFQHKDNDNPLRLVPRVY 120

QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 DB 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 QY 181 KLDELREGKASSAKORLKCSLQKFGERRAKANAVALSQRFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELREGKASSAKORLKCSLQKFGERRAKANAVALSQRFPKAEFAEVSKLVTDLTK 240
 QY 241 VHTECGHDLLECADDRADLAKYICENODSISSKLKECCERPLEKSHCIAVENDEMPA 300
 DB 241 VHTECGHDLLECADDRADLAKYICENODSISSKLKECCERPLEKSHCIAVENDEMPA 300
 QY 301 DLPSLADFEVSKDYCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
 DB 301 DLPSLADFEVSKDYCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDEKRPVLEEPPONLIKONCELEBOLGEYKFNALLVRYTKKVPYST 420
 DB 361 CAADPHCEYAKVDEKRPVLEEPPONLIKONCELEBOLGEYKFNALLVRYTKKVPYST 420
 QY 421 PTIVEFSNGLGKVSCKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTIVEFSNGLGKVSCKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKQTALVELVHKRPAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKQTALVELVHKRPAT 540
 QY 541 KEOLKAVDDPFAAFVEKCKKADDEKTCFAEGSKLVAASQALGL 585
 DB 541 KEOLKAVDDPFAAFVEKCKKADDEKTCFAEGSKLVAASQALGL 585

RESULT 5
 AAO20111
 ID AAO20111 standard; Protein: 585 AA.
 AC AAO20111;
 DT 06-AUG-2002 (first entry)
 XX HSA protein sequence related to the growth hormone protein.
 DE Serum albumin-growth hormone fusion protein; growth hormone;
 KW Down's syndrome.
 XX
 OS Unidentified.
 XX
 PN KR9076789-A.
 PD 15-OCT-1999.
 PF 25-JUN-1998; 98KR-0704914.
 PR 30-DEC-1995; 95GB-0026733.
 PR 19-DEC-1996; 96WO-GB03164.
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX WPI: 1997-363680/55.
 DR N-PSDB; AAK99568.
 XX
 PT Serum albumin-growth hormone fusion protein - useful to treat growth
 PT hormone related diseases, e.g. Down's syndrome
 XX
 PS Disclosure: Fig 6; 21pp: Korean.
 CC The invention relates to a serum albumin-growth hormone fusion protein -
 CC useful to treat growth hormone related diseases such as Down's syndrome.
 CC This sequence represents a HSA protein related to the serum albumin-
 CC growth hormone protein of the invention.
 XX
 SO Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 18; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRKFDGGEENFKALVLIARFAQYIQCCPFEDHVLVNEVEFAKTCVADSAAE 60
 DB 1 DAHSEVAHRKFDGGEENFKALVLIARFAQYIQCCPFEDHVLVNEVEFAKTCVADSAAE 60
 QY 61 NCDKSLHTLFGDKICTVATLRETYGEMADCCAKOPEPNECEFLQHKDNPMLPLVPREV 120
 DB 61 NCDKSLHTLFGDKICTVATLRETYGEMADCCAKOPEPNECEFLQHKDNPMLPLVPREV 120
 QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 DB 121 DVMCTAFHDNEETFLKKYLYEIAIRRHYPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 QY 181 KLDELREGKASSAKORLKCSLQKFGERRAKANAVALSQRFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELREGKASSAKORLKCSLQKFGERRAKANAVALSQRFPKAEFAEVSKLVTDLTK 240
 QY 241 VHTECGHDLLECADDRADLAKYICENODSISSKLKECCERPLEKSHCIAVENDEMPA 300
 DB 241 VHTECGHDLLECADDRADLAKYICENODSISSKLKECCERPLEKSHCIAVENDEMPA 300
 QY 301 DLPSLADFEVSKDYCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
 DB 301 DLPSLADFEVSKDYCKNYAEAKDVLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDEKRPVLEEPPONLIKONCELEBOLGEYKFNALLVRYTKKVPYST 420
 DB 361 CAADPHCEYAKVDEKRPVLEEPPONLIKONCELEBOLGEYKFNALLVRYTKKVPYST 420
 QY 421 PTIVEFSNGLGKVSCKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTIVEFSNGLGKVSCKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKQTALVELVHKRPAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKQTALVELVHKRPAT 540
 QY 541 KEOLKAVDDPFAAFVEKCKKADDEKTCFAEGSKLVAASQALGL 585
 DB 541 KEOLKAVDDPFAAFVEKCKKADDEKTCFAEGSKLVAASQALGL 585

RESULT 6
 AAY84873
 ID AAY84873 standard; protein: 585 AA.
 AC AAY84873;
 DT 08-AUG-2000 (first entry)
 XX Amino acid sequence of a human albumin protein.
 DE Human; albumin; ischemic state; serum protein; metal ion salt;
 KW perioperative ischemia; ischemia; myocardial infarction;
 KW progressive coronary artery disease.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 1 /note="optionally acetylated, and claimed under
 FT claim 56"
 XX
 PN WO200020840-A1.
 PD 13-APR-2000.
 PF 01-OCT-1999; 99WO-US22905.


```
Db      181 KILDELREGKASAKORLCKASLOKFGERRAKAMAVANLSORPFAEFAEVSCLVTDLTk 240
Qy      241 VHTCCGHDLECCADDDRADLAKYICENDOSTISSKLKECCCEPPLKESKICIAEVENDEMPA 300
Db      241 VHTCCGHDLECCADDDRADLAKYICENDOSTISSKLKECCCEPPLKESKICIAEVENDEMPA 300
Qy      301 DLPSIADPFVESKDYCKNAYEAKDYFLGMLPEYARRRHDYSVYLLRLAKTYETLEK 360
Db      301 DLPSIADPFVESKDYCKNAYEAKDYFLGMLPEYARRRHDYSVYLLRLAKTYETLEK 360
Qy      361 CAADPHCEYAKVDEKPELVEEPONLIKONCEPEQJGGEKFPONALLVRTKKVPOVST 420
Db      361 CAADPHCEYAKVDEKPELVEEPONLIKONCEPEQJGGEKFPONALLVRTKKVPOVST 420
Qy      421 PTIVEVSNNLKVSKCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTCTCTES 480
Db      421 PTIVEVSNNLKVSKCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTCTCTES 480
Qy      481 LVNRRPCSALEVEDETVYVPEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPRAT 540
Db      481 LVNRRPCSALEVEDETVYVPEFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPRAT 540
Qy      541 KEQLKAVDDPAAFYEKCCCKADDKETGFAEGSKYLAASQAALGL 585
Db      541 KEQLKAVDDPAAFYEKCCCKADDKETGFAEGSKYLAASQAALGL 585

RESULT 8
ABR79006 ID ABR79006 standard; Protein; 585 AA.
XX
XX ABR79006;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human mature albumin protein SEQ ID NO:18.
XX
KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytosstatic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;
KW type 1 diabetes mellitus; rheumatoid arthritis.
XX
OS Homo sapiens.
FH
FH Key location/Qualifiers
FT Domain 1..194
FT Domain 1..105
FT Domain /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..387
FT Domain /label= 2
FT Domain 195..291
FT Domain /label= subdomain
FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region 292..315
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT
```

```
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT Disulfide-bond /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567
XX
XX WO200179442-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US11850.
XX
XX 12-APR-2000; 2000US-229358P.
XX
XX 25-APR-2000; 2000US-199384P.
XX
XX 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Haseltine WA.
XX
XX WPI: 2001-611723/70.
XX
XX N-PSDB; ABR87288.
XX
XX New albumin fusion proteins, useful for treating diseases and disorders
XX such as cancer, comprise therapeutic protein fused to albumin
XX
XX Claim 1; Fig 11; 413pp; English.
XX
XX The present invention describes an albumin fusion protein (I) comprising
XX a therapeutic protein: X and (a fragment or variant of) albumin
XX comprising a the fully defined sequence in ABR79006 of 585 amino acids,
XX (where the fragment or variant has albumin or therapeutic protein: X
XX activity). (I) can have cytosstatic, anorectic, immunosuppressive,
XX antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX antibodies to C5, C242 and CD80 useful for treating various diseases
XX and disorders such as non-Hodgkin's lymphoma, cancer, obesity,
XX transplant rejection, type 1 diabetes mellitus, rheumatoid arthritis
XX and psoriasis. Fusing albumin to therapeutic proteins stabilises the
XX therapeutic protein, extends the shelf life and retains the in vitro or
XX in vivo biological activity. It also reduces the need to formulate
XX protein solutions with large excesses of carrier proteins to prevent
XX loss of therapeutic proteins due to factors such as binding to the
XX container. The fusion proteins are easily dispensed with a simple
XX formulation requiring minimal post storage manipulation. The fusion of
XX therapeutic proteins to albumin confers stability in aqueous or other
XX solution. The present sequence represents the mature human albumin (HA)
XX protein which is used in the exemplification of the present invention.
XX
XX Sequence 585 AA:
XX
XX Query Match 100.0%; Score 3103; DB 22; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 1e-254;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKEVAAHFKDLGEENFKALVLAFAQYLOOCPFEDHVKLVNEVTEFAKTVADESAE 60
Db 1 DAHKEVAAHFKDLGEENFKALVLAFAQYLOOCPFEDHVKLVNEVTEFAKTVADESAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGBMADCCAKQBERNECGLQHKDNPNIPLVRYREV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGBMADCCAKQBERNECGLQHKDNPNIPLVRYREV 120
Qy 121 DVMCTAFHNDNETPLKKYLYETIARRHPIYPAELLFFAKRYKAAPTECCQAADKAACLLP 180
```

```

Db      121 DVMCTAFHNDNEETFLKKYIETIARRHPYAPPELLFPARKYKAFTCCOADAADKACLLP 180
QY      181 KIDELRDEGKASSAKORLKCASLQKFGGERAFKMAVARLSORPPKAEFAVSKLVYDLTK 240
Db      181 KIDELRDEGKASSAKORLKCASLQKFGGERAFKMAVARLSORPPKAEFAVSKLVYDLTK 240
QY      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPILKESHCAIYENDMPA 300
Db      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPILKESHCAIYENDMPA 300
QY      301 DEPSLAADVEESKDVCKNYAEAKDVFGLGFLYETARRHPDYSVLLRLAKYETTLK 360
Db      301 DEPSLAADVEESKDVCKNYAEAKDVFGLGFLYETARRHPDYSVLLRLAKYETTLK 360
QY      361 CAADHDEGKAYKVFDEKFLVEBPONLIRKONCELFQOLGKYKONALLVRYTKKPVOST 420
Db      361 CAADHDEGKAYKVFDEKFLVEBPONLIRKONCELFQOLGKYKONALLVRYTKKPVOST 420
QY      421 PTLVEYSRLGKYGSCCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
Db      421 PTLVEYSRLGKYGSCCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
QY      481 LVNRRPCFSALEYDETYVPEKFNATFTFHADICTLSEKERQIKKOTALVELYKHKPKAT 540
Db      481 LVNRRPCFSALEYDETYVPEKFNATFTFHADICTLSEKERQIKKOTALVELYKHKPKAT 540
QY      541 KQOLKAVMDFAFVCKCKKADDKETCFABEGKKLVAAOALGL 585
Db      541 KQOLKAVMDFAFVCKCKKADDKETCFABEGKKLVAAOALGL 585

```

RESULT 9

AAEL3399 standard; Protein; 585 AA.

```

ID      AAEL3399 standard; Protein; 585 AA.
XX      AAEL3399;
XX      12-FEB-2002 (first entry)
DT      12-FEB-2002 (first entry)
XX      Human albumin (HA) protein.
XX      Human: albumin; HA: fusion protein; immune system disorder; syphilis;
KM      transplant rejection; blood related disorder; myocardial infarction;
KM      hyperproliferative disorder; acute myeloid leukemia; renal disorder;
KM      glomerulonephritis; cardiovascular disease; arthritis; rhinitis;
KM      respiratory disorder; neurological disease; Alzheimer's disease;
KM      endocrine disorder; pheochromocytoma; reproductive system disorder;
KM      measles; gastrointestinal disorder; irritable bowel syndrome; HIV;
KM      human immunodeficiency virus; wound healing; renal cell carcinoma;
KM      melanoma; gene therapy.
XX      Homo sapiens.
XX      Homo sapiens.
FH      Key
FH      Location/Qualifiers
FT      Domain
FT      /label= Loop-I
FT      /label= Loop-II
FT      /label= Loop-III
FT      /label= Loop-IV
FT      /label= Loop-V
FT      /label= Loop-VI
FT      /label= Loop-VII
FT      /label= Loop-VIII
FT      /label= Loop-IX
FT      Domain
FT      /label= Loop-IX

```

```

FT      Domain
FT      /label= Loop-X
FT      Domain
FT      /label= Loop-XI
FT      Domain
FT      /label= Loop-XII
PN      WO200179258-A1.
XX      25-OCT-2001.
PD      12-APR-2001; 2001WO-US12008.
PF      12-APR-2000; 2000US-129358P.
PR      25-APR-2000; 2000US-199384P.
PR      21-DEC-2000; 2000US-256931P.
XX      (HUMA-) HUMAN GENOME SCI INC.
PA      (PRIN-) PRINCIPAL PHARM CORP.
PI      Rosen CA, Sadeghi H, Prior CP, Turner AJ;
XX      WPI; 2001-602931/68.
DR      N-PDB; AAD22287.
XX      Albumin fusion proteins comprising a therapeutic protein and albumin,
PT      useful in the treating metastatic renal cell carcinoma, metastatic
PT      melanoma, malignant melanoma, renal cell carcinoma, HIV (human
XX      immunodeficiency virus) or infection -
XX      Claim 1; Fig 9; 325pp; English.
XX      The invention relates to albumin fusion proteins comprising therapeutic
XX      protein and human albumin (HA). The albumin fusion proteins are useful
XX      in the treatment, prevention, diagnosis, and/or detection of diseases,
XX      disorders such as immune system disorders (transplant rejection); blood
XX      related disorders (myocardial infarction); hyperproliferative disorders
XX      (childhood acute myeloid leukemia); renal disorder (glomerulonephritis);
XX      cardiovascular disorders (arrhythmias); respiratory disorders
XX      (non-allergic rhinitis); neurological diseases (Alzheimer's disease);
XX      endocrine disorders (pheochromocytoma); reproductive system disorders
XX      (syphilis); infectious diseases (measles); gastrointestinal disorders
XX      (irritable bowel syndrome) and wound healing. The albumin fusion
XX      proteins are also used in the treatment of metastatic renal cell
XX      carcinoma, metastatic melanoma, malignant melanoma and HIV (human
XX      immunodeficiency virus) infection. Nucleic acid encoding albumin fusion
XX      protein is useful in gene therapy. The present sequence is human
XX      albumin (HA) protein.
SO      Sequence 585 AA;

```

Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHSEVAHRFKDGLGENFKAVLTLFAQYLOCCPREDHVKLVNTEYFAKTCVADESA 60
Db      1 DAHSEVAHRFKDGLGENFKAVLTLFAQYLOCCPREDHVKLVNTEYFAKTCVADESA 60
QY      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAOEERNECFLOHDDNDNPLRLVRPEY 120
Db      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAOEERNECFLOHDDNDNPLRLVRPEY 120
QY      121 DVMCTAFHNDNEETFLKKYIETIARRHPYAPPELLFPARKYKAFTCCOADAADKACLLP 180
Db      121 DVMCTAFHNDNEETFLKKYIETIARRHPYAPPELLFPARKYKAFTCCOADAADKACLLP 180
QY      181 KIDELRDEGKASSAKORLKCASLQKFGGERAFKMAVARLSORPPKAEFAVSKLVYDLTK 240
Db      181 KIDELRDEGKASSAKORLKCASLQKFGGERAFKMAVARLSORPPKAEFAVSKLVYDLTK 240
QY      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPILKESHCAIYENDMPA 300
Db      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPILKESHCAIYENDMPA 300

```

Db	241	VHTECGHDDL	ECADRRADIAKYICENQDSISKLEKCECKPILLESKHCIAEVENDEMPA	300
Qy		301	DLPSTLAADFEVSEKDYCKNMYAEAKDYFLGMFLIYEYARRHBDYSVILLRLAKTYETTLKNC	360
Db		301	DLPSTLAADFEVSEKDYCKNMYAEAKDYFLGMFLIYEYARRHBDYSVILLRLAKTYETTLKNC	360
Qy		361	CAADPHCECYAKVFEPFEPKPLVVEEPQNLKQNCLELFPQLGEGYKQNLVLYRTKKVQVST	420
Db		361	CAADPHCECYAKVFEPFEPKPLVVEEPQNLKQNCLELFPQLGEGYKQNLVLYRTKKVQVST	420
Qy		421	PLIVSVSNLGVGSKCKCKHPAKRNPACADILSYVLNQLCYLHEKTPVSDRYTKCCTES	480
Db		421	PLIVSVSNLGVGSKCKCKHPAKRNPACADILSYVLNQLCYLHEKTPVSDRYTKCCTES	480
Qy		481	LVRNRPCTSALEVEDETVYYPKEFNAETFTPHADICTLSEKEROIKQATALVELYKHKPKAT	540
Db		481	LVRNRPCTSALEVEDETVYYPKEFNAETFTPHADICTLSEKEROIKQATALVELYKHKPKAT	540
Qy		541	KEQLKAVMDPFAAFYVEKCKKADDKETCFPEBEGKKLVAAQAALGL	585
Db		541	KEQLKAVMDPFAAFYVEKCKKADDKETCFPEBEGKKLVAAQAALGL	585
RESULT 10				
AAM52567				
ID			AAM52567 standard; Protein; 585 AA.	
XX			AAM52567;	
AC				
XX				
DT			05-FEB-2002 (first entry)	
XX				
DE			Mature human serum albumin.	
XX				
KM			Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiact;	
KM			nocrotropic; neuroprotective; gene therapy; immune disorder; wound healing;	
KM			hyperproliferative disorder; renal disorder; cardiovascular disorder;	
KM			respiratory disorder; neurological disease; endocrine disorder;	
KM			reproductive system disorder; infectious disease;	
KM			gastrointestinal disorder.	
XX				
OS			Homo sapiens.	
XX				
PN			WO200179444-A2.	
XX				
PD			25-OCT-2001.	
XX				
PF			12-APR-2001; 2001MO-US12013.	
XX				
PR			12-APR-2000; 2000US-229358P.	
PR			25-APR-2000; 2000US-199384P.	
XX			21-DEC-2000; 2000US-256931P.	
XX				
PA			(HUMA-) HUMAN GENOME SCI INC.	
XX				
PI			Rosen CA, Haseltine WA;	
XX				
DR			WPI; 2001-616755/71.	
XX			N-PSDB; ABA03057.	
PT			Albumin fusion proteins comprising a therapeutic protein and albumin,	
PT			useful in the treating immune system disorders (e.g. transplant	
PT			rejection), blood related disorders (e.g. myocardial infarction) and	
PT			hyperproliferative disorders -	
XX				
PS			Claim 1; Fig 15; 606pp; English.	
XX				
CC			The present invention relates to albumin fusion proteins, which comprise	
CC			a therapeutic protein and albumin. The present sequence is the protein	
CC			sequence for mature human serum albumin (HA), which was used to generate	
CC			the fusion proteins of the present invention. The albumin fusion proteins	
CC			are useful in the treatment, prevention, diagnosis, and/or detection of	
CC			diseases/disorders such as immune system disorders (e.g. transplant	
CC			rejection), blood related disorders (e.g. myocardial infarction),	

CC		hyperproliferative disorders (e.g. childhood acute myeloid leukemia), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arhythmias), respiratory disorders (e.g. non allergic rhinitis), neurological diseases (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis), infectious diseases (e.g. measles), gastrointestinal disorders (e.g. irritable bowel syndrome) and wound healing.
XX	Sequence	585 AA:
SQ	Match	100.0%; Score 3103; DB 22; Length 585;
	Best Local Similarity	100.0%; Pred. No. 1e-254;
	Matches	585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1	DAHSEVAHRRKDIGENFKALVLIAPAOYLQOCPEDHVKIYNVTFAKTCAVDSEAE 60
DB	1	DAHSEVAHRRKDIGSENFKAIVLIIAPAOYLQOCPPEEHVKIYNVTEFAKTCAVDSEAE 60
OY	61	NCDKSILHLFGDKICTAVTILRETTGEMADCCAKOEPERNECFLQHKDNPMLPRIVREV 120
DB	61	NCDKSILHLFGDKICTAVTLRRTYGEMADCCAKOEPERNECHLOHKDNPMLPRIVREV 120
OY	121	DVMCTAFHDNETFEKKLYEIAIRHPPIFYAPELLFFPKRYKAAETECQAADKAACLLP 180
DB	121	DVMCTAFHDNETFELKTYLIELARHPPIFYAPELLFFPKRYKAAETECQAADKAACLLP 180
OY	181	KIDELRDGCKASSAKORLKCSLOKFGERAFAKAWAVALRSOPFAEFRAEVSKLVTDLTk 240
DB	181	KIDELRDGCKASSAKORLKCSASIOKFGERAFKAWAVALRSOPFAEFRAEVSKLVTDLTk 240
OY	241	VHTCECHDDLEFCADRADLAIKYICENDSTISSKIECCERPLKSHCIAVENDEMPA 300
DB	241	VHTCECHDDLEFCADRADLAAYICENDSTISSKIECCERPLKSHCIAVENDEMPA 300
OY	301	DLPSIAADFVESKDVCKNYAEAKDYFLGMFTLEYEARHRDPDSVVLRLAKYTEETLEKC 360
DB	301	DLPSIAADFVESKDVCKNYAEAKDYFLGMFTLEYEARHRDPDSVVLRLAKYTEETLEKC 360
OY	361	CAAADPHCEYAKVFDEFPLVEEPPNLIKONCELFEOJGEYKFNALLVRYTKKPYOVST 420
DB	361	CAAADPHCEYAKVFDEFPLVEEPPNLIKONCELFEOJGEYKFNALLVRYTKKPYOVST 420
OY	421	PILYEVSURLGVNSCKCKHPAKRMPCAEIDLVSVLNOLCVLHKETPVSDVTCQPTES 480
DB	421	PILYEVSURLGVNSCKCKHPAKRMPCAEIDLVSVLNOLCVLHKETPVSDRYTCCTES 480
OY	481	IYNRRPCCSALEVDSTYYVPKEFNATFTFHADICLSKEROKIKOTALVELVKKRPKAT 540
DB	481	IYNRRPCCSALEVDSTYYVPKEFNATFTFHADICLSKEROKIKOTALVELVKKRPKAT 540
OY	541	KEQLAAVMDDEFAAVEKCCAKADKETCFEAESBKTLVAASOAALGL 585
DB	541	KEQLAAVMDDEFAAVEKCCAKADKETCFEAESBKTLVAASOAALGL 585
RESULT 11		
AAEI3129		
ID	AAEI3129	standard; Protein; 585 AA.
XX	AAEI3129;	
XX	28-JAN-2002	(first entry)
XX		
DE	Human albumin (HA).	
KW	Human; albumin; HA; fusion protein; therapeutic protein; vulnerable;	
KW	immune system disorder; transplant rejection; blood related disorder;	
KW	myocardial infarction; hyperproliferative disorder; glomerulonephritis;	
KW	childhood acute myeloid leukemia; cardiovascular disorder; arrythmia;	
KW	respiratory disorder; gene therapy; non-allergic rhinitis; noctopic;	
KW	neurological disorder; Alzheimer's disease; reproductive system disorder;	
KW	endocrine disorder; pheochromocytoma; infectious disease; antiarrhythmic;	
KW	measles; gastrointestinal disorders; irritable bowel syndrome; syphilis;	

KM wound healing; antiinflammatory; immunosuppressive; neuroprotective;
 KM cardiant; cytostatic; antileukaemic; antirheumatic; antimicrobial;
 KM renal disorder.
 OS Homo sapiens.
 XX Key
 FH Location/Qualifiers
 FT 54..61
 FT Domain /label= Loop_I
 FT 76..89
 FT Domain /label= Loop_II
 FT 92..100
 FT Domain /label= Loop_III
 FT 170..176
 FT Domain /label= Loop_IV
 FT 247..252
 FT Domain /label= Loop_V
 FT 266..277
 FT Domain /label= Loop_VI
 FT 280..288
 FT Domain /label= Loop_VII
 FT 362..368
 FT Domain /label= Loop_VIII
 FT 439..447
 FT Domain /label= Loop_IX
 FT 461..475
 FT Domain /label= Loop_X
 FT 478..486
 FT Domain /label= Loop_XI
 FT 560..566
 FT Domain /label= Loop_XII
 XX WO200179443-A2.
 XX 25-OCT-2001.
 XX 12-APR-2001; 2001WO-US11924.
 XX 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Haseltine WA;
 XX
 PI WPI; 2001-616754/71.
 DR N-PSDB; AAD21638.
 DR
 XX Albumin fusion proteins comprising a therapeutic protein and albumin;
 PT useful in the treating immune system disorders (e.g. transplant
 PT rejection), blood related disorders (e.g. myocardial infarction) and
 PT hyperproliferative disorders -
 XX
 PS Claim 1; Fig 9; 380pp; English.
 XX
 CC The invention relates to albumin fusion proteins comprising therapeutic
 CC protein and human albumin (HA). Therapeutic protein fused to albumin
 CC have an extended shelf-life. The albumin fusion proteins are useful in
 CC the treatment, prevention, diagnosis and/or detection of diseases,
 CC disorders such as immune system disorders (e.g. transplant rejection),
 CC blood related disorders (e.g. myocardial infarction), hyperproliferative
 CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
 CC respiratory disorders (e.g. non-allergic rhinitis), neurological
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
 CC pheochromocytoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome) and wound healing. Nucleic acids encoding
 CC albumin fusion protein is used in gene therapy. The present sequence
 CC is human albumin (HA) protein.
 CC
 XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No.1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFHFRKDLGEENFKALVLIAPAOYLQOCCPEPDHVKLVNVEYFAKTCVADSEAE 60
 DB 1 DAHSEVAFHFRKDLGEENFKALVLIAPAOYLQOCCPEPDHVKLVNVEYFAKTCVADSEAE 60
 QY 61 NCDSLHTLFQDKLCTVATLRETYGEMADCCAKOBERNCFYQHKDNDNLPRLVREPV 120
 DB 61 NCDSLHTLFQDKLCTVATLRETYGEMADCCAKOBERNCFYQHKDNDNLPRLVREPV 120
 QY 121 DVNCTAFHNDNEETFLKKYLYEIRRRHPYFAPRLFPFAKRYKAFTCCCAAKACCLP 180
 DB 121 DVNCTAFHNDNEETFLKKYLYEIRRRHPYFAPRLFPFAKRYKAFTCCCAAKACCLP 180
 QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAVRLSQRPKAFPAVSKLYVDLTK 240
 DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAVRLSQRPKAFPAVSKLYVDLTK 240
 QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLAECCERPLLEKSHCIAEYENDMPA 300
 DB 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLAECCERPLLEKSHCIAEYENDMPA 300
 QY 301 DLPSLADAFVESKDVCKNVAEAKDVLGMFLYETARRHPYSVLLRLAKYETTLK 360
 DB 301 DLPSLADAFVESKDVCKNVAEAKDVLGMFLYETARRHPYSVLLRLAKYETTLK 360
 QY 361 CAADPHECYAKYFDEKRPVEEPQNLKONCELFEOLGKRYKQNALVRYTKVPVST 420
 DB 361 CAADPHECYAKYFDEKRPVEEPQNLKONCELFEOLGKRYKQNALVRYTKVPVST 420
 QY 421 PTLEVSRNLGKVGSKCKHPEAKRMCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLEVSRNLGKVGSKCKHPEAKRMCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCELSLEVDETVVPRFNAEFTTFADICTISEKROTKKQATVLEVKHKRKAT 540
 DB 481 LVNRRPCELSLEVDETVVPRFNAEFTTFADICTISEKROTKKQATVLEVKHKRKAT 540
 QY 541 KEOLKAVMDPFAAFVEKCKKADKRETCFAEGKKLVASQAALGL 585
 DB 541 KEOLKAVMDPFAAFVEKCKKADKRETCFAEGKKLVASQAALGL 585

RESULT 12
 AAE12403
 ID AAE12403 standard; Protein; 585 AA.
 XX
 AC AAE12403;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human albumin (HA).
 XX
 XX Human; albumin; HA; immune system disorder; transplant rejection;
 KW blood related disorder; myocardial infarction; glomerulonephritis;
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;
 KW renal cell carcinoma; cardiovascular disorder; vulnary; melanoma;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;
 KW neurological disease; Alzheimer's disease; endocrine disorder; measles;
 KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;
 KW infectious disease; gastrointestinal disorder; wound healing; noctropic;
 KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;
 KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
 KW antiarrhythmic; antirheumatic; renal disorder; antimicrobial.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Domain Location/Qualifiers
 FT 54..61
 FT /label= Loop_I

PS Claim 1; Fig 1; 20pp; English.

XX The invention related to a method for testing cancer cells. The method is
CC useful for measuring human cancer cell proliferation, particularly for
CC determining the potential for inhibiting cancer cells proliferation using
CC albumin-derived peptides. The invention is also useful for drug screening
CC assays, as well as for evaluating biopsied tumours. The present sequence
CC is human serum albumin (HSA) related to the invention.

XX Sequence 585 AA:

Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGKFNKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESA 60
DB 1 DAHSEVAHFRKDLGKFNKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESA 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPNLRLVPRPY 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPNLRLVPRPY 120
QY 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPPELLFFAKRYKAFTCCOAAADKAACLP 180
DB 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPPELLFFAKRYKAFTCCOAAADKAACLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPKAEFAVSKLVYDTRK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPKAEFAVSKLVYDTRK 240
QY 241 VHTTECHGDLLECCADRADLAKTYICENODSISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTTECHGDLLECCADRADLAKTYICENODSISSKLECCCKPLLESHCIAEYENDMPA 300
QY 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADHCEYAKYDFEKKPLVEBPONLIKONCELFQDLGKRYKONALLVRYTKVQVST 420
DB 361 CAAADHCEYAKYDFEKKPLVEBPONLIKONCELFQDLGKRYKONALLVRYTKVQVST 420
QY 421 PTLVEFSRNIGKVGSCCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRTKCTTES 480
DB 421 PTLVEFSRNIGKVGSCCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRTKCTTES 480
QY 481 LVNRRCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKHKPKAT 540
DB 481 LVNRRCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKHKPKAT 540
QY 541 KEOLKAVMDPFAFVEKCKKADKKECTCFAEKGGKLVAAASQALGL 585
DB 541 KEOLKAVMDPFAFVEKCKKADKKECTCFAEKGGKLVAAASQALGL 585

RESULT 14

ID. ABG63321 standard; protein; 585 AA.

XX ABG63321;

DT 27-AUG-2002 (first entry)

XX Human serum albumin (HSA) protein.

XX Albumin fusion protein; therapeutic protein X; human albumin; HA;

KM human serum albumin; HSA; cancer; reproductive disorder;

KM digestive disorder; immune disorder; endocrine disorder;

KM haematopoietic disorder; neural disorder; connective disorder;

KM cytostatic; anti-infectivity; anti-inflammatory; anticancer;

KM immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;

KM neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;

KM osteopathic; antiarthritic.

XX Homo sapiens.

XX WO200177137-A1.

XX 18-OCT-2001.

XX 12-APR-2001; 2001WO-US11988.

XX 12-APR-2000; 2000US-229358P.

XX 25-APR-2000; 2000US-199384P.

XX 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2002-010886/01.

XX N-PSDB; ABK93280.

XX New fusion protein for treating disease e.g. diabetes comprises an

XX albumin fused to a therapeutic protein -

XX Claim 1; Fig 15; 2102pp; English.

CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). The present sequence represents HSA
CC (HA) protein.

XX Sequence 585 AA:

Query Match 100.0%; Score 3103; DB 23; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGKFNKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESA 60
DB 1 DAHSEVAHFRKDLGKFNKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESA 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPNLRLVPRPY 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPNLRLVPRPY 120
QY 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPPELLFFAKRYKAFTCCOAAADKAACLP 180
DB 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPPELLFFAKRYKAFTCCOAAADKAACLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPKAEFAVSKLVYDTRK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPKAEFAVSKLVYDTRK 240
QY 241 VHTTECHGDLLECCADRADLAKTYICENODSISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTTECHGDLLECCADRADLAKTYICENODSISSKLECCCKPLLESHCIAEYENDMPA 300
QY 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADHCEYAKYDFEKKPLVEBPONLIKONCELFQDLGKRYKONALLVRYTKVQVST 420
DB 361 CAAADHCEYAKYDFEKKPLVEBPONLIKONCELFQDLGKRYKONALLVRYTKVQVST 420

Db 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGGEYKFNQNALVRYTKVPOYST 420
Qy 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSRVTKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSRVTKCTES 480
Qy 481 LVNRRPCFSALEVDETVYPKFENAEFTFFHADICTLSKEKROIKQTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDETVYPKFENAEFTFFHADICTLSKEKROIKQTALVELVHKPKAT 540
Qy 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEKGKTLVAASQAALGI 585
Db 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEKGKTLVAASQAALGI 585

RESULT 15
ABJ00986
ID ABJ00986 standard; Protein: 585 AA.
XX
AC ABJ00986;
XX
DT 05-SEP-2002 (first entry)
XX
B lymphocyte stimulator protein binding peptide related protein.
DE
XX B lymphocyte stimulator protein binding protein: Blys; immune disease;
KW allergy; proliferative disease; infectious disease; arteriosclerosis;
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;
KW ischaemia; graft-versus-host disease; neurodegenerative disease;
KW immunosuppressive; nephrotropic; antirheumatic; antiathritic;
KW neutrophilic; cytosolic; immunostimulant; antitumor; anti-HIV;
KW antiaesthetic; antiallergic; thyromimetic; antianemic; haemostatic;
KW dermatological; antiinflammatory; cardiac; ophthalmological; uropathic;
KW antidiabetic; antithyroid; antidepressant; hepatotropic.
XX
OS Homo sapiens.
XX
PN MO200216411-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-US25850.
XX
PR 18-AUG-2000; 2000US-226700P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Beltzer JP, Potter DM, Fleming TL, Rosen CA;
XX
DR WPI; 2002-499775/53.
XX
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
PT administering B lymphocyte stimulator binding polypeptide
XX
PS Disclosure: Page 379-382; 387pp; English.
XX
XX The present invention relates to the treatment, prevention or
CC amelioration of a disease or disorder associated with: aberrant B
CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells
CC of haematopoietic origin; or proliferative disease; and reducing,
CC inhibiting or stimulating immunoglobulin production, B cell proliferation
CC and graft rejection involving administration of Blys binding polypeptide.
CC The Blys binding polypeptides are used in the treatment, prevention or
CC amelioration of diseases such as immune system diseases, proliferative
CC diseases, diseases of cells of hematopoietic origin, graft rejection,
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
CC neurodegenerative diseases. The present sequence is a protein described
CC in the invention.
XX
SQ Sequence 585 AA;

Query Match

100.0%; Score 3103; DB 23; Length 585;

Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRKRDGEENFKALVLAFAQYLOQCPFEDHVKLVNEVEFAKTCVADSAE 60
Db 1 DAHSEVAHRKRDGEENFKALVLAFAQYLOQCPFEDHVKLVNEVEFAKTCVADSAE 60
Qy 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOPEPNECFLOHKDNPMLPRVREV 120
Db 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOPEPNECFLOHKDNPMLPRVREV 120
Qy 121 DVNCTAHDNNEETFLKYLVEIARRHPYFAPELLFFAKRYKAFTBOCOADKAACLLP 180
Db 121 DVNCTAHDNNEETFLKYLVEIARRHPYFAPELLFFAKRYKAFTBOCOADKAACLLP 180
Qy 181 KLDELRDGKASAKORLKASLOKFGERRAKANAVARLSORFPAEFAEYSKLVTDLTJK 240
Db 181 KLDELRDGKASAKORLKASLOKFGERRAKANAVARLSORFPAEFAEYSKLVTDLTJK 240
Qy 241 VHTCCGHDDLECADDDRADLAKYICENODSISSKLKECEKPLLEKSHCIAVENDEMPA 300
Db 241 VHTCCGHDDLECADDDRADLAKYICENODSISSKLKECEKPLLEKSHCIAVENDEMPA 300
Qy 301 DLPSLADPFVSKDVCKNYAKADVFLGMFLYETARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLADPFVSKDVCKNYAKADVFLGMFLYETARRHPDYSVLLRLAKTYETTLK 360
Qy 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGGEYKFNQNALVRYTKVPOYST 420
Db 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGGEYKFNQNALVRYTKVPOYST 420
Qy 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSRVTKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSRVTKCTES 480
Qy 481 LVNRRPCFSALEVDETVYPKFENAEFTFFHADICTLSKEKROIKQTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDETVYPKFENAEFTFFHADICTLSKEKROIKQTALVELVHKPKAT 540
Qy 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEKGKTLVAASQAALGI 585
Db 541 KEOLKAVMDPFAAFVEKCKKADDKETCFAEKGKTLVAASQAALGI 585

Search completed: July 22, 2003, 11:43:43
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:43:03 ; Search time 18 Seconds
(without alignments)
956.245 Million cell updates/sec

Title: US-09-833-118-18
3103
Perfect score: 3103
Sequence: 1 DAHKEVAHREKDLGEENFK.....TCFAEGKKLVAAQALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
7: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	1	US-08-153-799-14 Sequence 14, Appl
2	3103	100.0	585	2	US-08-702-572-2 Sequence 2, Appl
3	3103	100.0	585	4	US-08-769-746-2 Sequence 2, Appl
4	3103	100.0	610	2	US-08-797-689-2 Sequence 2, Appl
5	3103	100.0	783	1	US-08-256-938-2 Sequence 2, Appl
6	3103	100.0	787	1	US-08-256-938-4 Sequence 4, Appl
7	3103	100.0	787	2	US-08-797-689-16 Sequence 16, Appl
8	3099	99.9	609	1	US-08-433-037-4 Sequence 3, Appl
9	3099	99.9	609	1	US-08-897-956A-2 Sequence 4, Appl
10	3099	99.9	609	5	PCT-US95-04075-3 Sequence 2, Appl
11	3095	99.7	978	4	US-08-897-956A-3 Sequence 3, Appl
12	3095	99.7	978	4	US-08-897-956A-3 Sequence 3, Appl
13	3093	99.7	585	1	US-08-448-196A-3 Sequence 3, Appl
14	3093	99.7	585	2	US-08-448-196A-3 Sequence 3, Appl
15	2458.5	79.2	583	1	US-08-448-196A-5 Sequence 1, Appl
16	2450.5	79.0	583	1	US-08-448-196A-5 Sequence 1, Appl
17	2432.5	78.4	583	1	US-08-448-196A-6 Sequence 4, Appl
18	2426	78.2	584	1	US-08-448-196A-6 Sequence 6, Appl
19	2389	77.0	582	1	US-08-134-638-1 Sequence 7, Appl
20	1249.5	40.3	609	1	US-08-222-619-4 Sequence 1, Appl
21	1249.5	40.3	609	5	PCT-US95-04075-4 Sequence 4, Appl
22	1206.5	38.9	590	2	US-09-186-723-2 Sequence 2, Appl
23	1206.5	38.9	590	4	US-08-505-012-5 Sequence 2, Appl
24	1206.5	38.9	590	4	US-08-505-012-5 Sequence 5, Appl
25	1206.5	38.9	590	4	US-09-186-949A-3 Sequence 3, Appl
26	1206.5	38.9	590	5	PCT-US96-00996-5 Sequence 5, Appl
27	1206.5	38.9	609	4	US-09-186-949A-2 Sequence 2, Appl

28	1164.5	37.5	579	1	US-08-448-196A-8	Sequence 8, Appl
29	1055	34.0	599	1	US-08-222-619-2	Sequence 2, Appl
30	1055	34.0	599	4	US-08-221-767-24	Sequence 24, Appl
31	1055	34.0	599	5	PCT-US95-04075-2	Sequence 2, Appl
32	926	29.8	393	2	US-08-377-309-7	Sequence 7, Appl
33	926	29.8	393	4	US-09-186-723-7	Sequence 7, Appl
34	926	29.8	393	4	US-08-505-012-10	Sequence 10, Appl
35	926	29.8	393	4	US-09-186-949A-8	Sequence 8, Appl
36	926	29.8	393	5	PCT-US96-00996-10	Sequence 10, Appl
37	777	25.0	324	4	US-08-505-012-12	Sequence 12, Appl
38	777	25.0	324	5	PCT-US96-00996-12	Sequence 12, Appl
39	777	25.0	325	2	US-08-377-309-8	Sequence 8, Appl
40	777	25.0	325	4	US-09-186-723-8	Sequence 8, Appl
41	777	25.0	325	4	US-08-505-012-11	Sequence 11, Appl
42	777	25.0	325	4	US-09-186-949A-9	Sequence 9, Appl
43	777	25.0	325	5	PCT-US96-00996-11	Sequence 11, Appl
44	747.5	24.1	590	1	US-08-448-196A-9	Sequence 9, Appl
45	684.5	22.1	389	2	US-08-377-309-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-153-799-14
Sequence 14, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear

```
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419 /note="Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note="Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 3103; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRPKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 1 DAHSEVAHRRPKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
QY 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
DB 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
QY 121 DVNCTAFHDNEETFLKLYEIAARRHPFYAPBELLFFAKRYKAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKLYEIAARRHPFYAPBELLFFAKRYKAFTCCQAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCSLOKFGBRAFKANAVARLSORFPKAEFAEYSKLVTDLT 240
DB 181 KLDELREGKASSAKORLKCSLOKFGBRAFKANAVARLSORFPKAEFAEYSKLVTDLT 240
QY 241 VHTCCGHDLLCEADDRAADLAKYICENODSISSKLKECEKPLEKSHCIAEVNDMPA 300
DB 241 VHTCCGHDLLCEADDRAADLAKYICENODSISSKLKECEKPLEKSHCIAEVNDMPA 300
QY 301 DLPSLADFEVSKVCKNYAKADVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADFEVSKVCKNYAKADVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
QY 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELEBOLGEYKFFONALLVRYTKKVPYST 420
DB 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELEBOLGEYKFFONALLVRYTKKVPYST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRYTKCTES 480
QY 481 LVNRRPCSALEVDVETVPKFFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
DB 481 LVNRRPCSALEVDVETVPKFFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEEGKRLVAASQAALGI 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEEGKRLVAASQAALGI 585

RESULT 2
US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon I.L.C.
STREET: 1020 First Avenue
```

```
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRPKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 1 DAHSEVAHRRPKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
QY 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
DB 61 NCDKSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
QY 121 DVNCTAFHDNEETFLKLYEIAARRHPFYAPBELLFFAKRYKAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKLYEIAARRHPFYAPBELLFFAKRYKAFTCCQAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCSLOKFGBRAFKANAVARLSORFPKAEFAEYSKLVTDLT 240
DB 181 KLDELREGKASSAKORLKCSLOKFGBRAFKANAVARLSORFPKAEFAEYSKLVTDLT 240
QY 241 VHTCCGHDLLCEADDRAADLAKYICENODSISSKLKECEKPLEKSHCIAEVNDMPA 300
DB 241 VHTCCGHDLLCEADDRAADLAKYICENODSISSKLKECEKPLEKSHCIAEVNDMPA 300
QY 301 DLPSLADFEVSKVCKNYAKADVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADFEVSKVCKNYAKADVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
QY 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELEBOLGEYKFFONALLVRYTKKVPYST 420
DB 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELEBOLGEYKFFONALLVRYTKKVPYST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHKEKTPVSDRYTKCTES 480
QY 481 LVNRRPCSALEVDVETVPKFFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
DB 481 LVNRRPCSALEVDVETVPKFFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
```

QY 541 KEOLKAYMDPFAAFVEKCKADKDKETCFABEGKTLVAASQAALGL 585
Db 541 KEOLKAYMDPFAAFVEKCKADKDKETCFABEGKTLVAASQAALGL 585

RESULT 3

US-08-769-746-2
Sequence 2, Application US/08769746
Patent No. 6274305
GENERAL INFORMATION:
APPLICANT: Sonenschein, Carlos
APPLICANT: Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,746
FILING DATE: 19-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRT-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHHRFKDGEENFKALVLI AFAOYLQCCPFEDHYKLVNEVEFAKTCVADDSAE 60
Db 1 DAHSEVAHHRFKDGEENFKALVLI AFAOYLQCCPFEDHYKLVNEVEFAKTCVADDSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLPLVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPRLPLVPEV 120
QY 121 DVNCTAHNDNEFTLKLYEIAARRHYFAPELLEFPAKRYKAAFTCCQAADKAACILP 180
Db 121 DVNCTAHNDNEFTLKLYEIAARRHYFAPELLEFPAKRYKAAFTCCQAADKAACILP 180
QY 181 KLDELREGKASSAKOKLKASLOKFGERRAFKAAVAVRLSORPKAEFAEYSKLVITLTK 240
Db 181 KLDELREGKASSAKOKLKASLOKFGERRAFKAAVAVRLSORPKAEFAEYSKLVITLTK 240
QY 241 VHEECGHGDLLECCADRADLAKYICENODSISSKLKECCCKPDLKSHCIAEVENDEMPA 300
Db 241 VHEECGHGDLLECCADRADLAKYICENODSISSKLKECCCKPDLKSHCIAEVENDEMPA 300
QY 301 DLPSLADPYESKVCNRYAKAVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLADPYESKVCNRYAKAVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHRECYAKVDEKFPVLEEPONLIKONCELEFEOLEGEKFORALLVRYTKKPOVST 420
Db 361 CAADPHRECYAKVDEKFPVLEEPONLIKONCELEFEOLEGEKFORALLVRYTKKPOVST 420
QY 421 PTVESVRNLGKVGSKCKKHEPAKRMPCAEYLSVNLQCVLHEKTPVSDRYTKCTES 480
Db 421 PTVESVRNLGKVGSKCKKHEPAKRMPCAEYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPFSALEVDETVPKFEFNAETFTPHADITLSEKEQIKOTALVELYKHKRYAT 540
Db 481 LVNRRPFSALEVDETVPKFEFNAETFTPHADITLSEKEQIKOTALVELYKHKRYAT 540
QY 541 KEOLKAYMDPFAAFVEKCKADKDKETCFABEGKTLVAASQAALGL 585
Db 541 KEOLKAYMDPFAAFVEKCKADKDKETCFABEGKTLVAASQAALGL 585

RESULT 4

US-08-797-689-2
Sequence 2, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guilton, Jean-Dominique
APPLICANT: Jung, Gerard
APPLICANT: Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 7e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNVEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNVEVTEFAKTCVADESAE 84
QY 61 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 120
DB 85 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 144
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFVAPPELLFFAKRYKAFTTECCQAADKACLLP 180
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFVAPPELLFFAKRYKAFTTECCQAADKACLLP 204
QY 181 KLDELRODGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 240
DB 205 KLDELRODGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 264
QY 241 VHTECCHDDLCECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHDDLCECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFESKDVCKNYAEKDVFLGMFLYEYARRPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFESKDVCKNYAEKDVFLGMFLYEYARRPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHBCYAKVDEFRPLVEEPONLIKONCELFEOLEGEYKFOALLVRYTKKPYQST 420
DB 385 CAADPHBCYAKVDEFRPLVEEPONLIKONCELFEOLEGEYKFOALLVRYTKKPYQST 444
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 504
QY 481 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKQOTALVELYKHKPRAT 540
DB 505 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKQOTALVELYKHKPRAT 564
QY 541 KEOLKAVMDPFAAFYEKCKKADDKETCFABEGRKLVAAASQAALGL 585
DB 565 KEOLKAVMDPFAAFYEKCKKADDKETCFABEGRKLVAAASQAALGL 609

RESULT 5
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne

REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNVEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNVEVTEFAKTCVADESAE 84
QY 61 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 120
DB 85 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 144
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFVAPPELLFFAKRYKAFTTECCQAADKACLLP 180
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFVAPPELLFFAKRYKAFTTECCQAADKACLLP 204
QY 181 KLDELRODGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 240
DB 205 KLDELRODGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 264
QY 241 VHTECCHDDLCECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHDDLCECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFESKDVCKNYAEKDVFLGMFLYEYARRPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFESKDVCKNYAEKDVFLGMFLYEYARRPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHBCYAKVDEFRPLVEEPONLIKONCELFEOLEGEYKFOALLVRYTKKPYQST 420
DB 385 CAADPHBCYAKVDEFRPLVEEPONLIKONCELFEOLEGEYKFOALLVRYTKKPYQST 444
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 504
QY 481 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKQOTALVELYKHKPRAT 540
DB 505 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKQOTALVELYKHKPRAT 564
QY 541 KEOLKAVMDPFAAFYEKCKKADDKETCFABEGRKLVAAASQAALGL 585
DB 565 KEOLKAVMDPFAAFYEKCKKADDKETCFABEGRKLVAAASQAALGL 609

RESULT 6
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA

```

? COUNTRY: USA
? ZIP: 19426
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: Macintosh
? OPERATING SYSTEM: System 7.1
? SOFTWARE: Word 5.0 (Patentin)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/256,938
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: FR 92/01065
? FILING DATE: 31-JAN-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Goodman, Rosanne
? REGISTRATION NUMBER: 32,534
? REFERENCE/DOCKET NUMBER: ST92007-US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (610) 454-3817
? TELEFAX: (610) 454-3808
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 787 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-256-938-4

```

```

Query Match 100.0%; Score 3103; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAKSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 60
DB 203 DAKHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 262
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 120
DB 263 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 322
QY 121 DVNCTAFHNDNEFLKLYLEIARRHPFYAPPELLFFAKRYKKAFTCCOADAACALLP 180
DB 323 DVNCTAFHNDNEFLKLYLEIARRHPFYAPPELLFFAKRYKKAFTCCOADAACALLP 382
QY 181 KLDELDEGKASSAKORLKCAASLOKFGERAFAKMAVAARLSQRPKAEFAVSKLYVDLTK 240
DB 383 KLDELDEGKASSAKORLKCAASLOKFGERAFAKMAVAARLSQRPKAEFAVSKLYVDLTK 442
QY 241 VHTCCGHDLECADRADLAKYICENODSISIKRBCCKRPLLEKSHCIAEYENDMPA 300
DB 443 VHTCCGHDLECADRADLAKYICENODSISIKRBCCKRPLLEKSHCIAEYENDMPA 502
QY 301 DLPSLADEVESKDVCKNYAEAKDVFLGMFLYEYARRHPYSVLLRLAKTYETTLK 360
DB 503 DLPSLADEVESKDVCKNYAEAKDVFLGMFLYEYARRHPYSVLLRLAKTYETTLK 562
QY 361 CAADPHECYAKVDEKFLVEEPQNLKONCELFEOLGSEYKFNALLVRYTKKVPVST 420
DB 563 CAADPHECYAKVDEKFLVEEPQNLKONCELFEOLGSEYKFNALLVRYTKKVPVST 622
QY 421 PTLVEVSRLNGKYGSCCKHPEAKRMPACADYLSVYLNOCVLHEKTPVSDRYTKCTES 480
DB 623 PTLVEVSRLNGKYGSCCKHPEAKRMPACADYLSVYLNOCVLHEKTPVSDRYTKCTES 682
QY 481 LVNRRPFSALVEDEVTPKPEFAETFTFHADICTLSEKEROIKKOTALVELYKHKPKAT 540
DB 683 LVNRRPFSALVEDEVTPKPEFAETFTFHADICTLSEKEROIKKOTALVELYKHKPKAT 742
QY 541 KEOLKAVMDPFAAFVEKCKADKKEFCFAEGKKLVAAASQALGL 585
DB 743 KEOLKAVMDPFAAFVEKCKADKKEFCFAEGKKLVAAASQALGL 787

```

```

RESULT 7
US-08-797-689-16
? Sequence 16, Application US/08797689
? Patent No. 5876969
? GENERAL INFORMATION:
? APPLICANT: Fleer, Reinhard
? APPLICANT: Fournier, Alain
? APPLICANT: Gultton, Jean-Dominique
? APPLICANT: Jung, Gerard
? TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
? PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
? TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
? NUMBER OF SEQUENCES: 36
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Rhone-Poulenc Rorer Inc.
? STREET: 500 Arcola Road, 3C43
? CITY: Collegeville
? STATE: PA
? COUNTRY: USA
? ZIP: 19426
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: Macintosh
? OPERATING SYSTEM: System 7.1
? SOFTWARE: Word 5.1 (Patentin)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/797,689
? FILING DATE: 31-JAN-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/256,927
? FILING DATE: 28-JUL-1994
? APPLICATION NUMBER: FR 92/01064
? FILING DATE: 31-JAN-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/FR93/00085
? FILING DATE: 28-JAN-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Smith Ph.D., Julie K.
? REGISTRATION NUMBER: P-38,619
? REFERENCE/DOCKET NUMBER: ST92006-US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (610) 454-3839
? TELEFAX: (610) 454-3808
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 787 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-797-689-16

```

```

Query Match 100.0%; Score 3103; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAKHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 60
DB 203 DAKHSEVAHRFKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADSAE 262
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 120
DB 263 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 322
QY 121 DVNCTAFHNDNEFLKLYLEIARRHPFYAPPELLFFAKRYKKAFTCCOADAACALLP 180
DB 323 DVNCTAFHNDNEFLKLYLEIARRHPFYAPPELLFFAKRYKKAFTCCOADAACALLP 382
QY 181 KLDELDEGKASSAKORLKCAASLOKFGERAFAKMAVAARLSQRPKAEFAVSKLYVDLTK 240
DB 383 KLDELDEGKASSAKORLKCAASLOKFGERAFAKMAVAARLSQRPKAEFAVSKLYVDLTK 442

```

QY 241 VHTCCGHDLLCEADDRADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
| | | | |
DB 443 VHTCCGHDLLCEADDRADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 502
QY 301 DLPSIADFEVSKDYCKNYAEAKDYFLGMPLEYARRRHDPYSVLLRLATYETTTLEKC 360
DB 503 DLPSIADFEVSKDYCKNYAEAKDYFLGMPLEYARRRHDPYSVLLRLATYETTTLEKC 562
QY 361 CAADPHCEYAKVDEFEKPLVEEPONLIKONCELFQEGYKFFONALLVRYTKKVPQVST 420
DB 563 CAADPHCEYAKVDEFEKPLVEEPONLIKONCELFQEGYKFFONALLVRYTKKVPQVST 622
QY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCTES 480
DB 623 PTLVEVSNLKGVSCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCTES 682
QY 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTLVELVHKRPAT 540
DB 683 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTLVELVHKRPAT 742
QY 541 KEOLKAVMDPFAAFVEKCCKADDKETCFAEBSGKLIYAASQAALGL 585
DB 743 KEOLKAVMDPFAAFVEKCCKADDKETCFAEBSGKLIYAASQAALGL 787

RESULT 8

US-08-222-619-3

: Sequence 3, Application US/08222619

: Patent No. 5652352

: GENERAL INFORMATION:

: APPLICANT: Lichenstein, Henri

: APPLICANT: Lyons, David

: APPLICANT: Wurfel, Mark

: APPLICANT: Wright, Samuel

: TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like

: TITLE OF INVENTION: Protein

: NUMBER OF SEQUENCES: 33

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Amgen Center, Patent Operations/RRC

: STREET: 1840 DeHavilland Drive

: CITY: Thousand Oaks

: STATE: California

: COUNTRY: U.S.

: ZIP: 91320-1789

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patentin Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/222,619

: FILING DATE:

: CLASSIFICATION: 435

: INFORMATION FOR SEQ ID NO: 3:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 609 amino acids

: TYPE: amino acid

: STRANDEDNESS: unknown

: TOPOLOGY: unknown

: MOLECULE TYPE: protein

US-08-222-619-3

Query Match 99.9%; Score 3099; DB 1; Length 609;

Best Local Similarity 99.8%; Pred. No. 1.7e-286;

Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDLGEENFKALVLAFAQYLQCCPFEDHVKLVNVEVFAATCVADESAE 60
| | | | |
DB 25 DAHSEVAHRRKDLGEENFKALVLAFAQYLQCCPFEDHVKLVNVEVFAATCVADESAE 84
| | | | |
QY 61 NCDSLSHLFGDKICTVATLTETTGEMADCCAKQEPERNECFLOHKDNPILRVLPREV 120
| | | | |

DB 85 NCDSLSHLFGDKICTVATLTETTGEMADCCAKQEPERNECFLOHKDNPILRVLPREV 144
QY 121 DVMCTAFHNDNETFLKTYELIARHHPYEAPELLFEAKRYKAAFTTECCQAADRAACLLP 180
DB 145 DVMCTAFHNDNETFLKTYELIARHHPYEAPELLFEAKRYKAAFTTECCQAADRAACLLP 204
QY 181 KLDELDEGKASAKQRLKASLQKFGERAKAMAVARLSORFPAEFAEVSKLVTDLTK 240
DB 205 KLDELDEGKASAKQRLKASLQKFGERAKAMAVARLSORFPAEFAEVSKLVTDLTK 264
QY 241 VHTCCGHDLLCEADDRADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCGHDLLCEADDRADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 324
QY 301 DLPSIADFEVSKDYCKNYAEAKDYFLGMPLEYARRRHDPYSVLLRLATYETTTLEKC 360
DB 325 DLPSIADFEVSKDYCKNYAEAKDYFLGMPLEYARRRHDPYSVLLRLATYETTTLEKC 384
QY 361 CAADPHCEYAKVDEFEKPLVEEPONLIKONCELFQEGYKFFONALLVRYTKKVPQVST 420
DB 385 CAADPHCEYAKVDEFEKPLVEEPONLIKONCELFQEGYKFFONALLVRYTKKVPQVST 444
QY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSNLKGVSCKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTLVELVHKRPAT 540
DB 505 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTLVELVHKRPAT 564
QY 541 KEOLKAVMDPFAAFVEKCCKADDKETCFAEBSGKLIYAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCCKADDKETCFAEBSGKLIYAASQAALGL 609

RESULT 9

US-08-433-037-4

: Sequence 4, Application US/08433037

: Patent No. 5707828

: GENERAL INFORMATION:

: APPLICANT: Sreekrishna, Kotlikanyadan

: APPLICANT: Brierley, Russell A.

: APPLICANT: Thill, Gregory P.

: APPLICANT: Tschopp, Juerg F.

: TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN

: TITLE OF INVENTION: PICHIA PASTORIS

: NUMBER OF SEQUENCES: 19

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Scully, Scott, Murphy & Presser

: STREET: 400 Garden City Plaza

: CITY: Garden City

: STATE: New York

: COUNTRY: U.S.A.

: ZIP: 11530-0299

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patentin Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/433,037

: FILING DATE: 03-MAY-1995

: CLASSIFICATION: 435

: ATTORNEY/AGENT INFORMATION:

: NAME: Digiglio, Frank S.

: REGISTRATION NUMBER: 31,346

: REFERENCE/DOCKET NUMBER: 9108Z

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (516) 742-4343

: TELEFAX: (516) 742-4366

: TELE: 230 901 SANS UR

: INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 1,7e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHRRKDGEEENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADSAE 60
DB 25 DAHSEVAHRRKDGEEENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADSAE 84
QY 61 NCCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNNECFLOHKDNDNPLPLVPRV 120
DB 85 NCCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNNECFLOHKDNDNPLPLVPRV 144
QY 121 DVNCTAFHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAADKAACLLP 180
DB 145 DVNCTAFHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAADKAACLLP 204
QY 181 KLDELDEGRKASSAKORLKCASLOKFGERAFAKAAVAVARLSQRPFAEFAEYSKLVTDLT 240
DB 205 KLDELDEGRKASSAKORLKCASLOKFGERAFAKAAVAVARLSQRPFAEFAEYSKLVTDLT 264
QY 241 VHTCCGDLLECCADDDADLAKTYICENODSISSKLKECCCKPILKSHCIAEVNDMPA 300
DB 265 VHTCCGDLLECCADDDADLAKTYICENODSISSKLKECCCKPILKSHCIAEVNDMPA 324
QY 301 DLPSLADPFVESKDVCKNVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLADPFVESKDVCKNVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRTKKVPVOST 420
DB 385 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRTKKVPVOST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 540
DB 505 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKTLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKTLVAASQAALGL 609
```

RESULT 10
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT FILING DATE: 1997-07-21
CURRENT APPLICATION NUMBER: US/08/897, 956A
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1,7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHRRKDGEEENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADSAE 60
DB 25 DAHSEVAHRRKDGEEENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADSAE 84
QY 61 NCCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNNECFLOHKDNDNPLPLVPRV 120
DB 85 NCCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNNECFLOHKDNDNPLPLVPRV 144
QY 121 DVNCTAFHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAADKAACLLP 180
DB 145 DVNCTAFHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAADKAACLLP 204
QY 181 KLDELDEGRKASSAKORLKCASLOKFGERAFAKAAVAVARLSQRPFAEFAEYSKLVTDLT 240
DB 205 KLDELDEGRKASSAKORLKCASLOKFGERAFAKAAVAVARLSQRPFAEFAEYSKLVTDLT 264
QY 241 VHTCCGDLLECCADDDADLAKTYICENODSISSKLKECCCKPILKSHCIAEVNDMPA 300
DB 265 VHTCCGDLLECCADDDADLAKTYICENODSISSKLKECCCKPILKSHCIAEVNDMPA 324
QY 301 DLPSLADPFVESKDVCKNVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLADPFVESKDVCKNVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRTKKVPVOST 420
DB 385 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRTKKVPVOST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 540
DB 505 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKTLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKTLVAASQAALGL 609
```

RESULT 11
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 DAHKEVAHRRKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEFEAKTCVADESAE 60
DB 25 DAHKEVAHRRKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEFEAKTCVADESAE 84
OY 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOPEERNECFLOKDDNPMLPRIVREPV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOPEERNECFLOKDDNPMLPRIVREPV 144
OY 121 DVMCTAFHDNETFLKKTLYELIARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 180
DB 145 DVMCTAFHDNETFLKKTLYELIARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 204
OY 181 KLDELDRGKASSAKORLKASLOKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 240
DB 205 KLDELDRGKASSAKORLKASLOKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 264
OY 241 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 300
DB 265 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 324
OY 301 DLPSTADPFVESKDVCKNYAEKADVFLGMFLYEVARRRHPDYSVLLRLAKYTEETLEKC 360
DB 325 DLPSTADPFVESKDVCKNYAEKADVFLGMFLYEVARRRHPDYSVLLRLAKYTEETLEKC 384
OY 361 CAADPHHCYAKVFEEFPLVEEPONLKONCELEKQGEYKFOALLVRYTKKPYOVST 420
DB 385 CAADPHHCYAKVFEEFPLVEEPONLKONCELEKQGEYKFOALLVRYTKKPYOVST 444
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRVTCCGES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRVTCCGES 504
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKHKPRAT 540
DB 505 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKHKPRAT 564
OY 541 KEQLAVMDPFAFVEKCKKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEQLAVMDPFAFVEKCKKADDKETCFABEGKRLVAASQAALGL 609
```

RESULT 12
US-08-897-956A-3
Sequence 3, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion polypeptides
FILE REFERENCE: 600-7244/GPA
CURRENT APPLICATION NUMBER: US/08/897, 956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 978
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.8%; Pred. No. 8.1e-286;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 DAHKEVAHRRKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEFEAKTCVADESAE 60
DB 212 DAHKEVAHRRKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNVEFEAKTCVADESAE 271
OY 61 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOPEERNECFLOKDDNPMLPRIVREPV 120
DB 272 NCDKSLHTLFGDKLCTVATLRETGYEMADCCAKOPEERNECFLOKDDNPMLPRIVREPV 331
OY 121 DVMCTAFHDNETFLKKTLYELIARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 180
DB 332 DVMCTAFHDNETFLKKTLYELIARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 391
OY 181 KLDELDRGKASSAKORLKASLOKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 240
DB 392 KLDELDRGKASSAKORLKASLOKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 451
OY 241 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 300
DB 452 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 511
OY 512 DLPSTADPFVESKDVCKNYAEKADVFLGMFLYEVARRRHPDYSVLLRLAKYTEETLEKC 571
DB 572 DLPSTADPFVESKDVCKNYAEKADVFLGMFLYEVARRRHPDYSVLLRLAKYTEETLEKC 631
OY 631 CAADPHHCYAKVFEEFPLVEEPONLKONCELEKQGEYKFOALLVRYTKKPYOVST 420
DB 652 CAADPHHCYAKVFEEFPLVEEPONLKONCELEKQGEYKFOALLVRYTKKPYOVST 611
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRVTCCGES 480
DB 632 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRVTCCGES 691
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKHKPRAT 540
DB 692 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKHKPRAT 751
OY 541 KEQLAVMDPFAFVEKCKKADDKETCFABEGKRLVAASQAALG 584
DB 752 KEQLAVMDPFAFVEKCKKADDKETCFABEGKRLVAASQAALG 795
```

RESULT 13
US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
Best Local Similarity 99.7%; Pred. No. 5,9e-286;

Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKLGEENFKALVLIAPQYLQCCPEDHVKLVNEYTEFAKTCVADESAAE 60
DB 1 DAKSEVAHRFKLGEENFKALVLIAPQYLQCCPEDHVKLVNEYTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNCFLOHKDNDNLPRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNCFLOHKDNDNLPRLVREY 120
QY 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
DB 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
QY 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
DB 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
QY 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
QY 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
QY 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
DB 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
QY 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
DB 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
QY 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
QY 541 KEOLKAVMDFAAFVEKCKADDKETCFEAEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDFAAFVEKCKADDKETCFEAEGKKLVAAASQAALGL 585

RESULT 14
US-08-984-176-1
Sequence 1, Application US/08984176

Patent No. 594609
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C
APPLICANT: HO, JOSEPH X
APPLICANT: RUKER, FLORIAN
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176
CURRENT APPLICATION NUMBER: US/08/984,176
CURRENT FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
Best Local Similarity 99.7%; Pred. No. 5,9e-286;

Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKLGEENFKALVLIAPQYLQCCPEDHVKLVNEYTEFAKTCVADESAAE 60
DB 1 DAKSEVAHRFKLGEENFKALVLIAPQYLQCCPEDHVKLVNEYTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNCFLOHKDNDNLPRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNCFLOHKDNDNLPRLVREY 120
QY 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
DB 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
QY 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
DB 121 DVNCTAFHNEETFLKYLIEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACLLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKGERAFKMAVAARLSQRPKAEFAVSKLVYDLTK 240
QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLESHCIAEYENDMPA 300
QY 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
QY 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
DB 301 DLPSLADEFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETLEKC 360
QY 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
DB 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
QY 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
DB 361 CAADHPECYAKVFDEKPLVEEPONLIKONCELFQLGKYEKFNMLLVYTKVPVST 420
QY 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEYSRNLGKYGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKHPRKAT 540
QY 541 KEOLKAVMDFAAFVEKCKADDKETCFEAEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDFAAFVEKCKADDKETCFEAEGKKLVAAASQAALGL 585

RESULT 15
US-08-448-196A-5
Sequence 5, Application US/08448196A

Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTRAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812

Job time : 20 secs

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/448,196A
;; FILING DATE: 23-MAY-1995
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROAD JR., ROBERT L.
;; REGISTRATION NUMBER: 18,757
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 205-544-0021
;; TELEFAX: 205-544-0258
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 583 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHEICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;

Best Local Similarity 75.8%; Pred. No. 1.5e-225;

Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;

OY 1 DAHSEVAHRRKDKGEENFKALVLAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DTHKSEIAHRRNDGEGEHFKGLVLAIFSQYLQCCPFEDHVKLVNEVTEFAKCAADESAE 60
OY 61 NCDSLHTLFEGDKICTATLTETGEMADCCAKOEPENECFLQKNDPNLPRLVREY 120
DB 61 NCDSLHTLFEGDKICTATLTETGEMADCCAKOEPENECFLQKNDPNLPRLVREY 119
OY 121 DVMCTAFHNEETFLKLYEIAARRHPYFAPPELLFPAKRYKAFTGCCQAADKACLLP 180
DB 121 DVMCTAFHNEETFLKLYEIAARRHPYFAPPELLFPAKRYKAFTGCCQAADKACLLP 179
OY 120 DAQCAAFQEDPDKGLGKLYIVARRHPYFPELLFPAEYKADTECCPADDKLACLIP 179
DB 120 DAQCAAFQEDPDKGLGKLYIVARRHPYFPELLFPAEYKADTECCPADDKLACLIP 179
OY 181 KLDELREDEGKASSAKQRLKCSLQKFGERAFAKAMAVARLSQFPKAEFAEYSKLVTDLTK 240
DB 181 KLDELREDEGKASSAKQRLKCSLQKFGERAFAKAMAVARLSQFPKAEFAEYSKLVTDLTK 239
OY 241 VHTCCCHDILLECADRDADLAKYICENDSISSKLKECCERPLEKSHCIAVENDEMPA 300
DB 241 VHTCCCHDILLECADRDADLAKYICENDSISSKLKECCERPLEKSHCIAVENDEMPA 299
OY 240 VHKBCCHDILLECADRDADLAKYICENDSISSKLKECCERPLEKSHCIAVENDEMPA 299
DB 240 VHKBCCHDILLECADRDADLAKYICENDSISSKLKECCERPLEKSHCIAVENDEMPA 299
OY 301 DLPSLADEFVSKDVCCKNYAKADVFGLMFLYEVARRHDPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLADEFVSKDVCCKNYAKADVFGLMFLYEVARRHDPDYSVLLRLAKTYETTLK 359
OY 300 DIPALADFAEDKELCKHYKAKADVFGLMFLYEVARRHDPDYSVLLRLAKTYETTLK 359
DB 300 DIPALADFAEDKELCKHYKAKADVFGLMFLYEVARRHDPDYSVLLRLAKTYETTLK 359
OY 361 CAADPHBCYAKVDFEKPVEEPONLTKONCELEFQLEGYKFNQALLVRYTKKVPQYST 420
DB 361 CAADPHBCYAKVDFEKPVEEPONLTKONCELEFQLEGYKFNQALLVRYTKKVPQYST 419
OY 360 CAADPHBCYAKVDFEKPVEEPONLTKONCELEFQLEGYKFNQALLVRYTKKVPQYST 419
DB 360 CAADPHBCYAKVDFEKPVEEPONLTKONCELEFQLEGYKFNQALLVRYTKKVPQYST 419
OY 421 PTLVEVSRLGKVSCKCKHPKAPKPCADYLSVNLQCVLHETKPYSDRYTKCTES 480
DB 421 PTLVEVSRLGKVSCKCKHPKAPKPCADYLSVNLQCVLHETKPYSDRYTKCTES 479
OY 420 PTLVEIGRTLGKVSCKCKHPKAPKPCADYLSVNLQCVLHETKPYSDRYTKCTES 479
DB 420 PTLVEIGRTLGKVSCKCKHPKAPKPCADYLSVNLQCVLHETKPYSDRYTKCTES 479
OY 481 LVNRRPCFSALEVDPTVYKPEFNATFTFHADICTLSKERQIKKOTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDPTVYKPEFNATFTFHADICTLSKERQIKKOTALVELVKHKPKAT 539
OY 480 LAERPCFSALEVDPTVYKPEFNATFTFHADICTLSKERQIKKOTALVELVKHKPKAT 539
DB 480 LAERPCFSALEVDPTVYKPEFNATFTFHADICTLSKERQIKKOTALVELVKHKPKAT 539
OY 541 KEOLKAVMDPFAAFVEKCCADKDETCFAEEGKKLVAAASQAL 583
DB 541 KEOLKAVMDPFAAFVEKCCADKDETCFAEEGKKLVAAASQAL 582
OY 540 KEOLKAVMDPFAAFVEKCCADKDETCFAEEGKKLVAAASQAL 582
DB 540 KEOLKAVMDPFAAFVEKCCADKDETCFAEEGKKLVAAASQAL 582

Search completed: July 22, 2003, 11:45:44

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:44:59 ; Search time 55 Seconds

(without alignments)
1263.172 Million cell updates/sec

Title: US-09-833-118-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHRFKDLGPEENK.....TCFAERGGKLVASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep2.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep3.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3103	100.0	585	Sequence 2, Appli
2	3103	100.0	585	Sequence 445, App
3	3103	100.0	585	Sequence 26, Appl
4	3103	100.0	585	Sequence 18, Appl
5	3103	100.0	585	Sequence 5, Appl
6	3103	100.0	604	Sequence 7, Appli
7	3103	100.0	609	Sequence 370, App
8	3103	100.0	609	Sequence 7, Appli
9	3103	100.0	610	Sequence 2, Appli
10	3103	100.0	610	Sequence 2, Appli
11	3103	100.0	610	Sequence 2, Appli
12	3103	100.0	610	Sequence 2, Appli
13	3103	100.0	610	Sequence 2, Appli
14	3103	100.0	610	Sequence 2, Appli
15	3103	100.0	651	Sequence 133, App
16	3103	100.0	652	Sequence 132, App

17	3103	100.0	653	US-10-153-064-131	Sequence 131, App
18	3103	100.0	656	US-10-153-064-130	Sequence 130, App
19	3103	100.0	676	US-10-153-064-127	Sequence 127, App
20	3103	100.0	676	US-10-153-064-129	Sequence 129, App
21	3103	100.0	677	US-10-153-064-125	Sequence 125, App
22	3103	100.0	680	US-10-153-064-123	Sequence 123, App
23	3103	100.0	787	US-09-984-186-16	Sequence 16, Appl
24	3103	100.0	787	US-10-237-667-16	Sequence 16, Appl
25	3103	100.0	787	US-10-237-708-16	Sequence 16, Appl
26	3103	100.0	787	US-10-237-866-16	Sequence 16, Appl
27	3103	100.0	787	US-10-237-871-16	Sequence 16, Appl
28	3103	100.0	787	US-10-237-624-16	Sequence 16, Appl
29	3103	100.0	788	US-10-073-118-26	Sequence 26, Appl
30	3092.5	99.7	652	US-10-153-064-96	Sequence 96, Appl
31	3092.5	99.7	652	US-10-153-064-99	Sequence 99, Appl
32	3092.5	99.7	652	US-10-153-064-105	Sequence 105, App
33	3092.5	99.7	660	US-10-153-064-90	Sequence 90, Appl
34	3092.5	99.7	660	US-10-153-064-93	Sequence 93, Appl
35	3092.5	99.7	676	US-10-153-064-95	Sequence 95, Appl
36	3092.5	99.7	676	US-10-153-064-98	Sequence 98, Appl
37	3092.5	99.7	676	US-10-153-064-104	Sequence 104, App
38	3092.5	99.7	684	US-10-153-064-92	Sequence 92, Appl
39	3092.5	99.7	1184	US-10-153-064-89	Sequence 89, Appl
40	3086.5	99.5	668	US-10-153-064-102	Sequence 102, App
41	3086.5	99.5	692	US-10-153-064-101	Sequence 101, App
42	2450.5	79.0	607	US-10-057-789-45	Sequence 45, Appl
43	2450.5	79.0	607	US-10-212-628-45	Sequence 45, Appl
44	2426	78.2	608	US-10-165-603-24	Sequence 24, Appl
45	2426	78.2	608	US-10-165-603-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-929-552-2
Sequence 2, Application US/09929552
Patent No. US20020123080A1
GENERAL INFORMATION:
APPLICANT: Sonmenscheln, Carlos
Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/929,552
FILING DATE: 14-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/769,746
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid

TOPOLGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOQCFEDHVKLVNEVTEFAATCVADSEAE 60
DB 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOQCFEDHVKLVNEVTEFAATCVADSEAE 60
QY 61 NCDSLHRLFGDKLCTVATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLYRREV 120
DB 61 NCDSLHRLFGDKLCTVATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLYRREV 120
QY 121 DVMCTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
DB 121 DVMCTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
QY 121 DVMTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
DB 121 DVMTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
QY 181 KLDELRDGKASAKORLKASLOKFGERRAKAMAVARLSQFPKAEAEVSKLVTDLTK 240
DB 181 KLDELRDGKASAKORLKASLOKFGERRAKAMAVARLSQFPKAEAEVSKLVTDLTK 240
QY 241 VHTBCHGDLLECADRADLAKYICENODSISSKIECCERPLLEKSHCIAEVENDEMPA 300
DB 241 VHTBCHGDLLECADRADLAKYICENODSISSKIECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
DB 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
QY 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
DB 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
QY 361 CAADPHBCYAKVDEFEPLVEEPONLTKONCELEEQGEYKFQNALLVRYTKKVPQYST 420
DB 361 CAADPHBCYAKVDEFEPLVEEPONLTKONCELEEQGEYKFQNALLVRYTKKVPQYST 420
QY 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTYSRVYKCTCES 480
DB 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTYSRVYKCTCES 480
QY 481 LVNRRPCSALEVEDETYPKPEFNAETFTFHADICTLSKEKQIKQTALVELVKKPKAT 540
DB 481 LVNRRPCSALEVEDETYPKPEFNAETFTFHADICTLSKEKQIKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCPAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCPAEEGKKLVAAASQAALGL 585

```

RESULT 2

US-09-932-613-445
Sequence 445, Application US/09932613
Publication No. US20030091565A1
GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
FILE REFERENCE: DXX-025.1 PCT: DXX-025.1 US
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: HomoSapiens
US-09-932-613-445

Query Match 100.0%; Score 3103; DB 12; Length 585;

Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOQCFEDHVKLVNEVTEFAATCVADSEAE 60
DB 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOQCFEDHVKLVNEVTEFAATCVADSEAE 60
QY 61 NCDSLHRLFGDKLCTVATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLYRREV 120
DB 61 NCDSLHRLFGDKLCTVATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLYRREV 120
QY 121 DVMCTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
DB 121 DVMCTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
QY 121 DVMTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
DB 121 DVMTAFHDNEETFLKYLLEYARHPYFVAPPELLFFAKRYKAATTECCQAADKACLLP 180
QY 181 KLDELRDGKASAKORLKASLOKFGERRAKAMAVARLSQFPKAEAEVSKLVTDLTK 240
DB 181 KLDELRDGKASAKORLKASLOKFGERRAKAMAVARLSQFPKAEAEVSKLVTDLTK 240
QY 241 VHTBCHGDLLECADRADLAKYICENODSISSKIECCERPLLEKSHCIAEVENDEMPA 300
DB 241 VHTBCHGDLLECADRADLAKYICENODSISSKIECCERPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
DB 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
QY 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
DB 301 DLPSLAADFEVSKDCKNYAEAKDVFGLMFLYETARRHPDYSVLLRLATYETTLTKC 360
QY 361 CAADPHBCYAKVDEFEPLVEEPONLTKONCELEEQGEYKFQNALLVRYTKKVPQYST 420
DB 361 CAADPHBCYAKVDEFEPLVEEPONLTKONCELEEQGEYKFQNALLVRYTKKVPQYST 420
QY 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTYSRVYKCTCES 480
DB 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTYSRVYKCTCES 480
QY 481 LVNRRPCSALEVEDETYPKPEFNAETFTFHADICTLSKEKQIKQTALVELVKKPKAT 540
DB 481 LVNRRPCSALEVEDETYPKPEFNAETFTFHADICTLSKEKQIKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCPAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCPAEEGKKLVAAASQAALGL 585

```

RESULT 3

US-09-984-010-26
Sequence 26, Application US/09984010
Publication No. US20030104578A1
GENERAL INFORMATION:

APPLICANT: Ballance, David James
TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984, 010
FILING DATE: 21-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/091, 873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEEENFKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHREKDLGEEENFKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDNDNPLRLVREY 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDNDNPLRLVREY 120
QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCASLOKGERAFKMAVAARLSQRPKAEFAVSKLYVDLTG 240
DB 181 KLDELREGKASSAKORLKCASLOKGERAFKMAVAARLSQRPKAEFAVSKLYVDLTG 240
QY 241 VHTECCHGDLLECADRADLAKYICENODSISSKLECKECPPLLEKSHCIAEVNDMPA 300
DB 241 VHTECCHGDLLECADRADLAKYICENODSISSKLECKECPPLLEKSHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVFDEKPLVEEPQNLKONCELFEGDGYEKKONALLVRYTKVPVST 420
DB 361 CAADPHECYAKVFDEKPLVEEPQNLKONCELFEGDGYEKKONALLVRYTKVPVST 420
QY 421 PTLVEVSRLNGKYGSKCKRPMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRLNGKYGSKCKRPMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVYKHKPKAT 540
DB 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVYKHKPKAT 540
QY 541 KEOLKAVMDDEFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDDEFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 585

RESULT 4
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; PRIOR APPLICATION NUMBER: 2001-04-12
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384

PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ. ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEEENFKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHREKDLGEEENFKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDNDNPLRLVREY 120
DB 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDNDNPLRLVREY 120
QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCASLOKGERAFKMAVAARLSQRPKAEFAVSKLYVDLTG 240
DB 181 KLDELREGKASSAKORLKCASLOKGERAFKMAVAARLSQRPKAEFAVSKLYVDLTG 240
QY 241 VHTECCHGDLLECADRADLAKYICENODSISSKLECKECPPLLEKSHCIAEVNDMPA 300
DB 241 VHTECCHGDLLECADRADLAKYICENODSISSKLECKECPPLLEKSHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVFDEKPLVEEPQNLKONCELFEGDGYEKKONALLVRYTKVPVST 420
DB 361 CAADPHECYAKVFDEKPLVEEPQNLKONCELFEGDGYEKKONALLVRYTKVPVST 420
QY 421 PTLVEVSRLNGKYGSKCKRPMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRLNGKYGSKCKRPMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVYKHKPKAT 540
DB 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVYKHKPKAT 540
QY 541 KEOLKAVMDDEFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDDEFAAFVEKCKADDKETCFABEGKKLVAAASQAALGL 585

RESULT 5
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ. ID NOS: 137
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 3103; DB 15; Length 585;
 Best Local Similarity 100.0%; Pred. No. 7,8e-271;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNEVEFAKTCVADSAE 60
 DB 1 DAHSEVAHRRKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNEVEFAKTCVADSAE 60
 QY 61 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPPLPRLVREY 120
 DB 61 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPPLPRLVREY 120
 QY 121 DVMTAFHNDNEETFLKLYEIAARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 DB 121 DVMTAFHNDNEETFLKLYEIAARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 QY 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
 DB 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
 QY 241 VHTCCGHDLLCEADDRLADLAKYICENDSISSKLKECEKPLEKSHCIAVENDEMPA 300
 DB 241 VHTCCGHDLLCEADDRLADLAKYICENDSISSKLKECEKPLEKSHCIAVENDEMPA 300
 QY 301 DLPSLAADFESKDVCKNYAEAKDVFLGMFLYEARHDPDYSVLLRLAKTYETTLK 360
 DB 301 DLPSLAADFESKDVCKNYAEAKDVFLGMFLYEARHDPDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELEQIGERYKFOALLVRYTKKPOVST 420
 DB 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELEQIGERYKFOALLVRYTKKPOVST 420
 QY 421 PTIVEVSNTLKVSGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTIVEVSNTLKVSGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKEKQIKQTALVELVKKRPAT 540
 DB 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKEKQIKQTALVELVKKRPAT 540
 QY 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGSKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGSKLVAAASQAALGL 585

RESULT 6

US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1

GENERAL INFORMATION:

APPLICANT: Ballance, David James
 TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 AND SERUM ALBUMIN

NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:

ADDRESSSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
 STREET: 1300 I Street, NW
 CITY: Washington
 STATE: DC

COUNTRY: USA
 ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/984,010
 FILING DATE: 21-May-2002

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/091,873
 FILING DATE: 25-JUN-1998

APPLICATION NUMBER: PCT/G996/03164
 FILING DATE: 19-DEC-1996
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:

LENGTH: 604 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>

TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-984-010-7
 Query Match 100.0%; Score 3103; DB 12; Length 604;
 Best Local Similarity 100.0%; Pred. No. 8,2e-271;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNEVEFAKTCVADSAE 60
 DB 20 DAHSEVAHRRKDLGEENFKALVLIAPAYIQCCPFEDHVKLVNEVEFAKTCVADSAE 79
 QY 61 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPPLPRLVREY 120
 DB 80 NCDSLHFLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPPLPRLVREY 139
 QY 121 DVMTAFHNDNEETFLKLYEIAARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 DB 140 DVMTAFHNDNEETFLKLYEIAARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 199
 QY 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
 DB 200 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 259
 QY 241 VHTCCGHDLLCEADDRLADLAKYICENDSISSKLKECEKPLEKSHCIAVENDEMPA 300
 DB 260 VHTCCGHDLLCEADDRLADLAKYICENDSISSKLKECEKPLEKSHCIAVENDEMPA 319
 QY 301 DLPSLAADFESKDVCKNYAEAKDVFLGMFLYEARHDPDYSVLLRLAKTYETTLK 360
 DB 320 DLPSLAADFESKDVCKNYAEAKDVFLGMFLYEARHDPDYSVLLRLAKTYETTLK 379
 QY 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELEQIGERYKFOALLVRYTKKPOVST 420
 DB 380 CAADPHCEYAKVDEFPPLVEEPONLIKONCELEQIGERYKFOALLVRYTKKPOVST 439
 QY 421 PTIVEVSNTLKVSGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 440 PTIVEVSNTLKVSGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 499
 QY 481 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKEKQIKQTALVELVKKRPAT 540
 DB 500 LVNRRPCSALEVEDETVPKPEFNAETFFHADICTLSKEKQIKQTALVELVKKRPAT 559
 QY 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGSKLVAAASQAALGL 585
 DB 560 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGSKLVAAASQAALGL 604

RESULT 7

US-09-919-039-370
 ; Sequence 370, Application US/09919039
 ; Publication No. US20030108871A1

GENERAL INFORMATION:

APPLICANT: Kaser, Matthew R.
 TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US

CURRENT FILING DATE: 2002-09-09

PRIOR APPLICATION NUMBER: 60/222,113
 PRIOR FILING DATE: 2000-07-28
 NUMBER OF SEQ ID NOS: 401
 SOFTWARE: PERL Program

```

; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 088957CD1
US-09-919-039-370

```

```

Query Match      100.0%; Score 3103; DB 12; Length 609;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPREDHYKLVNTEYTERAKTVADESAB 60
DB 25 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPREDHYKLVNTEYTERAKTVADESAB 84
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHKKDDNPMLPRLVPEY 120
DB 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHKKDDNPMLPRLVPEY 144
QY 121 DVWCTAFHNEEFELKKYIETARRHPYAPPELLFPARKYKAFTCCOADAADKACLP 180
DB 145 DVWCTAFHNEEFELKKYIETARRHPYAPPELLFPARKYKAFTCCOADAADKACLP 204
QY 181 KIDELDEGKASSAKORLKCASLOKGERAFKMAVAVARLSQRPKAFPAEVSCLVTDLR 240
DB 205 KIDELDEGKASSAKORLKCASLOKGERAFKMAVAVARLSQRPKAFPAEVSCLVTDLR 264
QY 241 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCEKPLEKSHCIAEYENDEMPA 300
DB 265 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCEKPLEKSHCIAEYENDEMPA 324
QY 301 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVYLLRLAKYETTLK 360
DB 325 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVYLLRLAKYETTLK 384
QY 361 CAADHPECYAKYFDEFKPLVEBPONLIONCELFQGLGEYKRONALLVRYTKKPVQVST 420
DB 385 CAADHPECYAKYFDEFKPLVEBPONLIONCELFQGLGEYKRONALLVRYTKKPVQVST 444
QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDEYVPEKFAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
DB 505 LVNRRPCFSALVEDEYVPEKFAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 564
QY 541 KEOLKAVMDFAAFVFEKCKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEOLKAVMDFAAFVFEKCKADDKETCFABEGKRLVAASQAALGL 609

```

```

RESULT 8
US-10-153-064-7
; Sequence 7, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

```

```

Query Match      100.0%; Score 3103; DB 15; Length 609;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPREDHYKLVNTEYTERAKTVADESAB 60
DB 25 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPREDHYKLVNTEYTERAKTVADESAB 84
QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHKKDDNPMLPRLVPEY 120
DB 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHKKDDNPMLPRLVPEY 144
QY 121 DVWCTAFHNEEFELKKYIETARRHPYAPPELLFPARKYKAFTCCOADAADKACLP 180
DB 145 DVWCTAFHNEEFELKKYIETARRHPYAPPELLFPARKYKAFTCCOADAADKACLP 204
QY 181 KIDELDEGKASSAKORLKCASLOKGERAFKMAVAVARLSQRPKAFPAEVSCLVTDLR 240
DB 205 KIDELDEGKASSAKORLKCASLOKGERAFKMAVAVARLSQRPKAFPAEVSCLVTDLR 264
QY 241 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCEKPLEKSHCIAEYENDEMPA 300
DB 265 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCEKPLEKSHCIAEYENDEMPA 324
QY 301 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVYLLRLAKYETTLK 360
DB 325 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVYLLRLAKYETTLK 384
QY 361 CAADHPECYAKYFDEFKPLVEBPONLIONCELFQGLGEYKRONALLVRYTKKPVQVST 420
DB 385 CAADHPECYAKYFDEFKPLVEBPONLIONCELFQGLGEYKRONALLVRYTKKPVQVST 444
QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDEYVPEKFAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
DB 505 LVNRRPCFSALVEDEYVPEKFAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 564
QY 541 KEOLKAVMDFAAFVFEKCKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEOLKAVMDFAAFVFEKCKADDKETCFABEGKRLVAASQAALGL 609

```

```

RESULT 9
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 3103; DB 11; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DAHSEVAHRRKDGEEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
25 DAHSEVAHRRKDGEEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 84
61 NCDSLHRTLFSGDKLCTVAATLTRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 120
85 NCDSLHRTLFSGDKLCTVAATLTRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 144
121 DVMTAFHNDNETFLKLYELIARRHPYFYAPPELLFFAKRYKAFTTECCQAADRAACLLP 180
145 DVMTAFHNDNETFLKLYELIARRHPYFYAPPELLFFAKRYKAFTTECCQAADRAACLLP 204
181 KLDELRDGKASSAKORIKCASLOKFGERRAFKANAVARLSORFPAEAEVSKLVTDLTK 240
205 KLDELRDGKASSAKORIKCASLOKFGERRAFKANAVARLSORFPAEAEVSKLVTDLTK 264
241 VHTCCGHDLLECCADDRADLAKYICENDSISIKLECCERPLEKSHCIAEVENDMPA 300
265 VHTCCGHDLLECCADDRADLAKYICENDSISIKLECCERPLEKSHCIAEVENDMPA 324
301 DLPSLAADFVSKYCKNYAKADYFLGMFLYERARRPDYSVYLLRLATYETTLTKC 360
325 DLPSLAADFVSKYCKNYAKADYFLGMFLYERARRPDYSVYLLRLATYETTLTKC 384
361 CAADPHHCYAKVDFEFLVEEPONLIKQNCLEPEQGEVKKFONALLVRYTKKPVYST 420
385 CAADPHHCYAKVDFEFLVEEPONLIKQNCLEPEQGEVKKFONALLVRYTKKPVYST 444
421 PTLVEVSNTLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 480
445 PTLVEVSNTLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 504
481 LVNRRPCSALEVDVETYPKREFNAETFTFHADICTLSKERQIKKQTLVLYVKKRPAT 540
505 LVNRRPCSALEVDVETYPKREFNAETFTFHADICTLSKERQIKKQTLVLYVKKRPAT 564
541 KEOLKAVMDPFAAFYEKCCCKADDEKTCFAEGGKILVAASOALGL 585
565 KEOLKAVMDPFAAFYEKCCCKADDEKTCFAEGGKILVAASOALGL 609

RESULT 10
US-10-237-667-2
Sequence 2, Application US/10237667
Publication No. US20030022308A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES.
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3543
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-667-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DAHSEVAHRRKDGEEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
25 DAHSEVAHRRKDGEEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 84
61 NCDSLHRTLFSGDKLCTVAATLTRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 120
85 NCDSLHRTLFSGDKLCTVAATLTRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREY 144
121 DVMTAFHNDNETFLKLYELIARRHPYFYAPPELLFFAKRYKAFTTECCQAADRAACLLP 180
145 DVMTAFHNDNETFLKLYELIARRHPYFYAPPELLFFAKRYKAFTTECCQAADRAACLLP 204
181 KLDELRDGKASSAKORIKCASLOKFGERRAFKANAVARLSORFPAEAEVSKLVTDLTK 240

Db 205 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 264
 QY 241 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 Db 265 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADVESKDKCNKNAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
 Db 325 DLPSLAADVESKDKCNKNAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 384
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCEFLQGEYKFNALLVRYTKYQVST 420
 Db 385 CAADPHCEYAKVFDEKFLVEEPONLIKONCEFLQGEYKFNALLVRYTKYQVST 444
 QY 421 PTLVEYSRLKGVGSCCKHPRKMPKCAEDYLSVNLQCVLHETPVS DRYTKCTES 480
 Db 445 PTLVEYSRLKGVGSCCKHPRKMPKCAEDYLSVNLQCVLHETPVS DRYTKCTES 504
 QY 481 LVNRRPCFSALVEDETYVKEFNAETFTPHADICTLSEKEROIKKOTALVELYKHKPKAT 540
 Db 505 LVNRRPCFSALVEDETYVKEFNAETFTPHADICTLSEKEROIKKOTALVELYKHKPKAT 564
 QY 541 KEQLKAVMDFAAFVEKCKADDKETCFAEBSKLIVAASQAALGL 585
 Db 565 KEQLKAVMDFAAFVEKCKADDKETCFAEBSKLIVAASQAALGL 609

RESULT 11
 US-10-237-708-2
 ; Sequence 2, Application US/10237708
 ; Publication No. US20030036170A1
 ; GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Guitton, Jean-Dominique
 Jung, Gerard
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/237,708
 FILING DATE: 10-Sep-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION NUMBER:
 APPLICATION NUMBER: US/08/797,689
 FILING DATE: 31-Jan-1997
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-Jul-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-Jan-1992

ATTORNEY/AGENT INFORMATION:
 NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808

; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 610 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-237-708-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
 Best Local Similarity 100.0%; Pred. No. 8.3e-271;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSFAHFRFDIGENFKALVLTAFADYLOCCPEHDHVKLVNTEEFKACVADSEAE 60
 Db 25 DAHKSFAHFRFDIGENFKALVLTAFADYLOCCPEHDHVKLVNTEEFKACVADSEAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDNPMLRLVPRPV 120
 Db 85 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDNPMLRLVPRPV 144
 QY 121 DVWCTAFHNEETFLKLYEYIARRHPYYPABELLFFAKRYAAFTCCQAADKAACLIP 180
 Db 145 DVWCTAFHNEETFLKLYEYIARRHPYYPABELLFFAKRYAAFTCCQAADKAACLIP 204
 QY 181 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 240
 Db 205 KIDELRDEGKASSAKORLKASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 264
 QY 241 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 Db 265 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADVESKDKCNKNAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
 Db 325 DLPSLAADVESKDKCNKNAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETLEKC 384
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCEFLQGEYKFNALLVRYTKYQVST 420
 Db 385 CAADPHCEYAKVFDEKFLVEEPONLIKONCEFLQGEYKFNALLVRYTKYQVST 444
 QY 421 PTLVEYSRLKGVGSCCKHPRKMPKCAEDYLSVNLQCVLHETPVS DRYTKCTES 480
 Db 445 PTLVEYSRLKGVGSCCKHPRKMPKCAEDYLSVNLQCVLHETPVS DRYTKCTES 504
 QY 481 LVNRRPCFSALVEDETYVKEFNAETFTPHADICTLSEKEROIKKOTALVELYKHKPKAT 540
 Db 505 LVNRRPCFSALVEDETYVKEFNAETFTPHADICTLSEKEROIKKOTALVELYKHKPKAT 564
 QY 541 KEQLKAVMDFAAFVEKCKADDKETCFAEBSKLIVAASQAALGL 585
 Db 565 KEQLKAVMDFAAFVEKCKADDKETCFAEBSKLIVAASQAALGL 609

RESULT 12
 US-10-237-866-2
 ; Sequence 2, Application US/10237866
 ; Publication No. US20030036171A1
 ; GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Guitton, Jean-Dominique
 Jung, Gerard
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRPKDGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRRPKDGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 120
DB 85 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 144
QY 121 DVNCTAFHNDNEETFLKTYEIAIRRHPIFYAPPELLFEAKRYKAAFTCCQAADRAACILP 180
DB 145 DVNCTAFHNDNEETFLKTYEIAIRRHPIFYAPPELLFEAKRYKAAFTCCQAADRAACILP 204
QY 181 KLDELROGKASSAKQRLKASLQKFGRAKAMAVANLSORFPAEAEVSKLVTDTLK 240
DB 205 KLDELROGKASSAKQRLKASLQKFGRAKAMAVANLSORFPAEAEVSKLVTDTLK 264
QY 241 VTECCGHDLLCECADDRLAKYICENDSISSKLKECCCKPILLESKICIAEVENDEMPA 300
DB 265 VTECCGHDLLCECADDRLAKYICENDSISSKLKECCCKPILLESKICIAEVENDEMPA 324
QY 301 DLPSLAADVEESKDVCKNVAEADVFLGMLPEYARRRHDYSVLLRLAATYETTELK 360
DB 325 DLPSLAADVEESKDVCKNVAEADVFLGMLPEYARRRHDYSVLLRLAATYETTELK 384
QY 361 CAADPHHCYAKVDEFPPLVEEPONLIKONCELFEOJGEEKFQONALLVRTTKVPOYST 420
DB 385 CAADPHHCYAKVDEFPPLVEEPONLIKONCELFEOJGEEKFQONALLVRTTKVPOYST 444
QY 421 PTLVEVSNLGKVSCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCES 480
DB 445 PTLVEVSNLGKVSCKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSERQIKQOTALVELVKKRPAT 540

DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSERQIKQOTALVELVKKRPAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKETCFAGEGKTLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKETCFAGEGKTLVAASQAALGL 609

RESULT 13
US-10-237-871-2
Sequence 2, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3643
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-871-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRPKDGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRRPKDGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 120

```

Db      85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLVPRPV 144
QY      121 DVMCTAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 180
Db      145 DVMCTAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 204
QY      181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPKAFPAEVSCLVTDLRK 240
Db      205 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPKAFPAEVSCLVTDLRK 264
QY      241 VHTCCGHDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
Db      265 VHTCCGHDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSTLAADVFESKDKVKNYAEAKDVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 360
Db      325 DLPSTLAADVFESKDKVKNYAEAKDVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 384
QY      361 CAADPHCEYAKYFDEKFLVEEPONLIKONCELFQOLGEYKQNALVRYTKRPQVST 420
Db      385 CAADPHCEYAKYFDEKFLVEEPONLIKONCELFQOLGEYKQNALVRYTKRPQVST 444
QY      421 PTLVEYSRNLGVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCCTES 480
Db      445 PTLVEYSRNLGVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCCTES 504
QY      481 LVNRRCFSALVEDEYVYKPEFNAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
Db      505 LVNRRCFSALVEDEYVYKPEFNAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 564
QY      541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 585
Db      565 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 609

```

RESULT 14 US-10-237-624-2

; Sequence 2, Application US/10237624
; Publication No. US20030082747A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fourtner, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patricia

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentlin)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237, 624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797, 689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256, 927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085

```

; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-624-2

```

Query Match 100.0%; Score 3103; DB 15; Length 610;

Best Local Similarity 100.0%; Pred. No. 8.3e-271;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHSEVAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 60
Db      25 DAHSEVAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 84
QY      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLVPRPV 120
Db      85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLVPRPV 144
QY      121 DVMCTAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 180
Db      145 DVMCTAFHNEETFLKKYLYETARRHPYAPPELLFFAKRYKAAPTECCOADAACCLIP 204
QY      181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPKAFPAEVSCLVTDLRK 240
Db      205 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPKAFPAEVSCLVTDLRK 264
QY      241 VHTCCGHDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
Db      265 VHTCCGHDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSTLAADVFESKDKVKNYAEAKDVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 360
Db      325 DLPSTLAADVFESKDKVKNYAEAKDVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 384
QY      361 CAADPHCEYAKYFDEKFLVEEPONLIKONCELFQOLGEYKQNALVRYTKRPQVST 420
Db      385 CAADPHCEYAKYFDEKFLVEEPONLIKONCELFQOLGEYKQNALVRYTKRPQVST 444
QY      421 PTLVEYSRNLGVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCCTES 480
Db      445 PTLVEYSRNLGVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRYTKCCTES 504
QY      481 LVNRRCFSALVEDEYVYKPEFNAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
Db      505 LVNRRCFSALVEDEYVYKPEFNAETFTFHADICTISEKEROIKKOTALVELYKHKPKAT 564
QY      541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 585
Db      565 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKKLVAAASQALGL 609

```

RESULT 15

US-10-153-064-133

; Sequence 133, Application US/10153064

; Publication No. US20020142814A1

GENERAL INFORMATION:

APPLICANT: Bell et al.

TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

FILE REFERENCE: PF556

CURRENT APPLICATION NUMBER: US/10/153, 064

CURRENT FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 3103; DB 15; Length 651;
Best Local Similarity 100.0%; Pred. No. 9,1e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRRFKDGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRRFKDGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVPREV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVPREV 186
QY 121 DVNCTAFHDNEETFLKKLYLEIARRHRYFYAPELLFFAKRYKAFTGCCQAADKAACILP 180
DB 187 DVNCTAFHDNEETFLKKLYLEIARRHRYFYAPELLFFAKRYKAFTGCCQAADKAACILP 246
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKANAVALSQRPPKAEPAEVSKLVTDLT 240
DB 247 KLDELDEGKASSAKORLKCASLOKFGERAFAKANAVALSQRPPKAEPAEVSKLVTDLT 306
QY 241 VHTCCGHDLLCEADDRADLAKYICENODSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 307 VHTCCGHDLLCEADDRADLAKYICENODSISSKLKECCCKPILLEKSHCIAEVENDEMPA 366
QY 301 DLPSLADPVESSKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
DB 367 DLPSLADPVESSKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK 426
QY 361 CAADPHECYAKVDEEKPVEEPONL IKONCELEFQLEGEYKFQNALLVRYTKKVPQVST 420
DB 427 CAADPHECYAKVDEEKPVEEPONL IKONCELEFQLEGEYKFQNALLVRYTKKVPQVST 486
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHCKTPVSDRYTKCCTES 480
DB 487 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHCKTPVSDRYTKCCTES 546
QY 481 LVNRRPCFSALEVEDETVVPKEFNATFTFHADICTLSSEKROIKKOTALVELVKKHPRAT 540
DB 547 LVNRRPCFSALEVEDETVVPKEFNATFTFHADICTLSSEKROIKKOTALVELVKKHPRAT 606
QY 541 KEQIKAVMDPFAFVEKCKKADKKEFCFAEGKKLVAAASQAALGL 585
DB 607 KEQIKAVMDPFAFVEKCKKADKKEFCFAEGKKLVAAASQAALGL 651

Search completed: July 22, 2003, 11:53:57
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:42:49 ; Search time 22 Seconds
(Without alignments)
2556.302 Million cell updates/sec

Title: US-09-833-118-18

Perfect score: 3103

Sequence: 1 DAHSEVAREHFKDLEGENFK.....TCFAEGKRLVAASQALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

283224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ABHUS
2	2942	94.8	609	2	A47391
3	2620	84.4	608	2	S57632
4	2475.5	79.8	607	1	ABHOS
5	2446.5	78.8	607	1	ABHOS
6	2432.5	78.4	607	1	ABSHS
7	2426	78.2	608	1	ABRTS
8	2411.5	77.7	605	1	ABPGS
9	2387	76.9	609	2	JC5838
10	1861	60.0	453	2	A05139
11	1557.5	50.2	615	1	ABCHS
12	1253.5	40.4	609	2	JC4258
13	1249.5	40.3	609	1	PPCO
14	1242.5	40.0	609	1	PPHU
15	1205	38.8	607	1	ABXL72
16	1181.5	38.1	265	2	I46986
17	1175.5	37.9	608	1	ABXL68
18	1084	34.9	605	1	FPMS
19	1067	34.4	611	1	FPRT
20	1055	34.0	599	1	A54906
21	928.5	29.9	614	2	A55917
22	928	29.9	608	2	A53195
23	747.5	24.1	608	1	ABON61
24	742.5	23.9	608	1	ABON62
25	699	22.5	382	2	A37253
26	440.5	14.2	1423	2	S27941
27	386	12.4	474	1	VYHDD
28	385	12.4	476	1	VYRTD
29	372	12.0	472	1	A35327

30	184	5.9	1819	2	A71928	cag island protein
31	184	5.9	1927	2	G64585	cag pathogenicity
32	141.5	4.6	1560	2	T30282	calcium-binding pr
33	137.5	4.4	1348	2	AG2558	hypothetical prote
34	134.5	4.3	1004	2	JC2221	major surface glyco
35	134	4.3	1083	2	JC2300	cell surface glyco
36	134	4.3	1780	2	T11722	hypothetical prote
37	133.5	4.3	1390	2	S51364	sperm tail-specifi
38	132.5	4.3	1070	2	T06733	kinesin homolog F2
39	131	4.2	1076	2	JC2217	major surface glyco
40	130	4.2	1175	2	D35815	myosin heavy chain
41	129	4.2	1017	2	PC4035	cell-cycle-depende
42	128.5	4.1	1051	2	JC4091	glycoprotein A - p
43	128	4.1	1005	2	A64465	hypothetical prote
44	128	4.1	3225	2	I52300	giantin - human
45	126.5	4.1	1927	2	A59236	embryonic muscle m

ALIGNMENTS

RESULT 1

ABHOS serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

C:Species: Homo sapiens (man)

C:Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text, change 17-Mar-2000

C:Accession: A93743; A93936; I59427; I59286; I59313; G01747; S53144; A91420; S06422;

R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najjarian, R.C.; See

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia c

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K', 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J00078; GB:U00132; GB:U00133; NID:928591; PIDN:CA

R:Dugalczyk, A.; Law, S.W.; Dennison, O.E.

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936

A:Molecule type: mRNA

A:Residues: 1-120, 'G', 122-609 <DUG>

A:Cross-references: EMBL:V00494; NID:928589; PIDN:CA23753.1; PID:928590

R:Umano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions a

A:Reference number: I39427; MUID:86140099; PMID:2419329

A:Accession: I39427

A:Molecule type: not shown

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <URA>

A:Cross-references: GB:M13075; NID:9178330; PIDN:AAA51688.1; PID:9553173

R:Watkins, S.; Madison, J.; Galliano, M.; Minichotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian faml

A:Reference number: I59286; MUID:94181575; PMID:8134387

A:Accession: I59286

A:Molecule type: DNA

A:Status: translated from GB/EMBL/DDBJ

A:Residues: 282-290, 'KSRPDIQ' <MAT>

A:Cross-references: GB:S69192; NID:9546032; PIDN:AB30282.1; PID:9546033

A>Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia

R:Madison, J.; Galliano, M.; Watkins, S.; Minichotti, L.; Porta, F.; Rossi, A.; Putna

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carbox

A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313

A:Molecule type: DNA

A:Status: translated from GB/EMBL/DDBJ

A:Residues: 589-590, 'ALPRRVNLLQVRLP' <MAD>

A:Cross-references: GB:S70799; NID:9547231; PIDN:AB31177.1; PID:9547232

A:Note: this frame-shift variant is designated albumin Bazzano; four additional variants
 R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
 Submitted to the EMBL Data Library, March 1995
 A:Reference number: G08292
 A:Accession: G01747
 A>Status: translated from GB/EMBL/DBDUB
 A:Molecule type: mRNA
 A:Residues: 1-120, 'G', 122-455 <MEN>
 A:Cross-references: EMBL:U25961; NID:q763428; PIDN:AAA64922.1; PID:q763431
 R:Idgewood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
 Biochem. J. 308, 321-325, 1995
 A:Title: Entoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
 A:Reference number: S55314; MUID:95275251; PMID:7755581
 A:Accession: S55314
 A:Molecule type: protein
 A:Residues: 19-27 <LED>
 R:Mejoun, B.; Moravsek, L.; Kostka, V.
 FEBS Lett. 58, 134-137, 1975
 A:Title: Complete amino acid sequence of human serum albumin.
 A:Reference number: A91420; MUID:76187907; PMID:1225573
 A:Accession: A91420
 A:Molecule type: protein
 A:Residues: 25-117, 'E', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
 R:Roehr, U.; Spittler, G.; Tripler, D.
 Justus Liebig's Ann. Chem. 9, 881-884, 1988
 A:Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
 A:Reference number: S06422
 A:Note: this paper is in German, with an English abstract
 A:Accession: S06422
 A:Molecule type: protein
 A:Residues: 25-48 <RO>
 R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
 Arch. Biochem. Biophys. 305, 595-599, 1993
 A:Title: Mass spectrometric identification of modifications to human serum albumin treat
 A:Reference number: S36882; MUID:93384321; PMID:8373198
 A:Accession: S36882
 A:Molecule type: protein
 A:Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>
 R:Kausler, E.; Spittler, G.
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
 A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmo
 A:Reference number: S17599; MUID:92126241; PMID:1772598
 A:Accession: S17599
 A:Molecule type: protein
 A:Residues: 25-54, 354-357, 431-447 <KAU>
 A:Note: 49-Leu was also found
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: A45800
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986
 A:Title: The amino acid sequence of kineinsin, a novel peptide isolated from pepsin-tre
 A:Reference number: A03239; MUID:86242180; PMID:3087352
 A:Accession: A03239
 A:Molecule type: protein
 A:Residues: 166-173, 'U' <MOG>
 R:Galliano, M.; Minichotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.
 A:Reference number: A38255; MUID:91062352; PMID:2247440
 A:Accession: C38255
 A:Molecule type: protein
 A:Residues: 76-111 <GAL1>
 A:Accession: B38255
 A:Molecule type: protein
 A:Residues: 82-105, 'K', 107-110 <GAL2>
 A:Note: this variant is designated albumin Vibo Valentia
 A:Accession: A38255
 A:Molecule type: protein

A:Residues: 76-83, 'K', 85-106 <GAL3>
 A:Note: this variant is designated albumin Torino
 R:Minichotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
 Eur. J. Biochem. 214, 437-444, 1993
 A:Title: The structural characterization and bilirubin-binding properties of albumin
 A:Reference number: S33298; MUID:93292504; PMID:8513793
 A:Accession: S33298
 A:Molecule type: protein
 A:Residues: 255-263, 'E', 265-281 <MIN1>
 A:Note: this variant is designated albumin Herborn
 R:Minichotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Por
 Biochim. Biophys. Acta 1119, 237-238, 1992
 A:Title: Two allolobins with identical electrophoretic mobility are produced by dif
 A:Reference number: S21078; MUID:92190239; PMID:1347703
 A:Accession: S21078
 A:Molecule type: protein
 A:Residues: 354-356, 'K', 358-378 <MIN2>
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is report
 R:He, X.M.; Carter, D.C.
 Nature 358, 209-215, 1992
 A:Title: Atomic structure and chemistry of human serum albumin.
 A:Reference number: A46756; MUID:92334427; PMID:1630489
 A:Contents: annotation: X-ray crystallography, 2.8 angstroms
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp. 23-
 A:Reference number: A94442
 A:Contents: annotation: three-dimensional structure and disulfide bonds
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Mejoun, B.
 Collect. Czech. Chem. Commun. 42, 564-579, 1977
 A:Title: Disulfide bonds in human serum albumin.
 A:Reference number: A90930
 A:Contents: annotation: disulfide bonds
 R:Jacobsen, C.
 Biochem. J. 171, 453-459, 1978
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bindi
 A:Reference number: A90289; MUID:78186630; PMID:656055
 A:Contents: annotation: bilirubin-binding site
 R:Peterson, T.; Reed, R.G.
 In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoblom, I., eds., 11-
 A:Title: Serum albumin: conformation and active sites.
 A:Reference number: A94408
 A:Contents: annotation: binding sites
 R:Harper, M.E.; Dugalczyk, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein ge
 A:Reference number: A90028; MUID:83279982; PMID:6192711
 A:Contents: annotation: gene position
 R:Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic ac
 A:Reference number: A46755; MUID:76257808; PMID:955075
 A:Contents: annotation
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic a
 R:Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
 FEBS Lett. 298, 266-268, 1992
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phos
 A:Reference number: A56294; MUID:92183881; PMID:1544460
 A:Contents: annotation
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; 1
 atase activity
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (we
 C:Comment: A large number of variants of human serum albumin have been described.
 C:Genetics:
 A:Gene: GDB:ALB
 A:Cross-references: GDB:118990; OMTM:103600
 A:Map position: 4q11-4q13
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyri
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status experimental <PRO>
 F:25-609/Product: serum albumin #status experimental <PT>
 F:29-202/Domain: serum albumin repeat homology <SA1>

F:166-174/Product: kinetensin #status experimental <RIP>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA2>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,148-193,192-201,224-270,265-277,289-303,302-313,340-385,384-393,4
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 7,4e-198;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRFDLGEENFKALVLIAPAOYLQOCPEDEHVKLVNTEFEAKTCVADESA 60
 |||||
 DB 25 DAHSEVAHFRFDLGEENFKALVLIAPAOYLQOCPEDEHVKLVNTEFEAKTCVADESA 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDNDPMLPRVREPV 120
 |||||
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDNDPMLPRVREPV 144
 QY 121 DVMCTAFHNEETFLKKYLETARHPYFAPPELLFFAKRYKAFTGCCOADKAACLP 180
 |||||
 DB 145 DVMCTAFHNEETFLKKYLETARHPYFAPPELLFFAKRYKAFTGCCOADKAACLP 204
 QY 181 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 240
 |||||
 DB 205 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 264
 QY 241 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 300
 |||||
 DB 265 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 324
 QY 301 DPLSLAADVVEESKDVCKNAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLK 360
 |||||
 DB 325 DPLSLAADVVEESKDVCKNAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLK 384
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 420
 |||||
 DB 385 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 444
 QY 421 PTLVEYSRLIGVSGCKCHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 |||||
 DB 445 PTLVEYSRLIGVSGCKCHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 504
 QY 481 LVNRRPCFSALEDETYVYKERNAEFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
 |||||
 DB 505 LVNRRPCFSALEDETYVYKERNAEFTFHADICTISEKEROIKKOTALVELYKHKPKAT 564
 QY 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKTLVAASQALGL 585
 |||||
 DB 565 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKTLVAASQALGL 609

RESULT 2

serum albumin precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A47391
 R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwalet, J.; Putnam, F.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliary
 A:Reference number: A47391; MUID:93211971; PMID:8460152
 A:Contents: B/B homozygote
 A:Accession: A47391
 A:Status: preliminary
 A:Molecule type: mRNA; protein
 A:Residues: 1-600 <MAT>
 A:Cross-references: GB:M90463; NID:9342294; PIDN:AAA36906.1; PID:9342295
 A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBI:128280, NCBI:128281)
 C:Superfamily: serum albumin; serum albumin repeat homology
 F:21-194/Domain: serum albumin repeat homology <SA1>
 F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2942; DB 2; Length 600;
 Best Local Similarity 93.5%; Pred. No. 3.3e-187;
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRFDLGEENFKALVLIAPAOYLQOCPEDEHVKLVNTEFEAKTCVADESA 60
 |||||
 DB 17 DAHSEVAHFRFDLGEENFKALVLIAPAOYLQOCPEDEHVKLVNTEFEAKTCVADESA 76
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDNDPMLPRVREPV 120
 |||||
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECFLOHKDNDPMLPRVREPV 136
 QY 121 DVMCTAFHNEETFLKKYLETARHPYFAPPELLFFAKRYKAFTGCCOADKAACLP 180
 |||||
 DB 137 DVMCTAFHNEETFLKKYLETARHPYFAPPELLFFAKRYKAFTGCCOADKAACLP 196
 QY 181 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 240
 |||||
 DB 197 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 256
 QY 241 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 300
 |||||
 DB 257 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 316
 QY 301 DPLSLAADVVEESKDVCKNAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLK 360
 |||||
 DB 317 DPLSLAADVVEESKDVCKNAEAKDVLGMFLVEYARHPDYSVLLRLAKTYETTLK 376
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 420
 |||||
 DB 377 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 436
 QY 421 PTLVEYSRLIGVSGCKCHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
 |||||
 DB 437 PTLVEYSRLIGVSGCKCHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 496
 QY 481 LVNRRPCFSALEDETYVYKERNAEFTFHADICTISEKEROIKKOTALVELYKHKPKAT 540
 |||||
 DB 497 LVNRRPCFSALEDETYVYKERNAEFTFHADICTISEKEROIKKOTALVELYKHKPKAT 556
 QY 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKTLVAASQALGL 583
 |||||
 DB 557 KEOLKAVMDFAAFVEKCKKADDEKTCFAEEGKTLVAASQALGL 599

RESULT 3

serum albumin precursor - cat
 C:Species: Felis silvestris catus (domestic cat)
 C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
 C:Accession: J04660; S57632
 R:Hiiger, C.; Grigioni, F.; Hentges, F.
 Gene 169, 295-296, 1996
 A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
 A:Reference number: J04660; MUID:96194824; PMID:8647469
 A:Accession: J04660
 A:Molecule type: mRNA
 A:Residues: 1-608 <H12>
 A:Cross-references: EMBL:84842; NID:9886484; PIDN:CAA59279.1; PID:9886485
 A:Experimental source: liver
 A:Comment: This protein is the major protein component in plasma. It functions as a m
 eln has 35 conserved cysteine residues.
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: liver; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRP>
 F:25-608/Product: serum albumin #status predicted <MAT>
 F:29-202/Domain: serum albumin repeat homology <SA1>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>

[illegible]

A:Residues: 1-32 <MAG>
 R:Hisah, J.C.; Lin, F.P.; Tam, M.F.
 Anal. Biochem. 170, 1-8, 1988
 A:Title: Electrophoretic transfer of glass-fiber filter from an analytical isoelectrofocusing gel
 A:Reference number: A60808; MUID:88267456; PMID:3389500
 A:Accession: B60808
 A:Molecule type: protein
 A:Residues: 25-41 <HSI>
 R:Strawich, E.; Glincher, M.J.
 Eur. J. Biochem. 191, 47-56, 1990
 A:Title: Tooth enamelins identified mainly as serum proteins. Major 'enamelin' is albumin
 A:Reference number: S10780; MUID:90336641; PMID:2379503
 A:Accession: S10780
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-57, 59-64 <SNR>
 R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: D45800
 A:Molecule type: protein
 A:Residues: 163-172 <CAR>
 R:Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
 J. Biol. Chem. 262, 5968-5973, 1987
 A:Title: Structure of a biologically active neurotensin-related peptide obtained from pe
 A:Reference number: A26693; MUID:87194805; PMID:2437111
 A:Accession: A26693
 A:Molecule type: protein
 A:Residues: 165-172, 'L', <CA2>
 R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.
 Biochem. J. 191, 867-868, 1980
 A:Title: Sequence of residues 400-403 of bovine serum albumin.
 A:Reference number: A90309; MUID:82023364; PMID:7283978
 A:Accession: A90309
 A:Molecule type: protein
 A:Residues: 402-433 <REE>
 R:Brown, J.R.
 Fed. Proc. 34, 591, 1975
 A:Title: Structure of bovine serum albumin.
 A:Reference number: A91458
 A:Accession: A91458
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'H'
 R:Brown, J.R.
 Submitted to the Atlas, April 1975
 A:Reference number: A94551
 A:Accession: A94551
 A:Molecule type: protein
 A:Residues: 190-195 <BR2>
 R:Brown, J.R.
 Fed. Proc. 33, 1389, 1974
 A:Reference number: A91457
 A:Contents: annotation; disulfide bonds
 R:Werlen, R.C.; Olford, R.E.; Rose, K.
 Biochem. J. 302, 907-911, 1994
 A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC
 A:Reference number: S55232; MUID:95031935; PMID:7945219
 A:Accession: S55232
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 529-536; 569-572 <MER>
 C:Superfamily: serum albumin; copper binding; duplication; plasma
 C:Keywords: carrier protein; copper binding; duplication; plasma
 F:1-18/Domain: signal sequence #status experimental <PRO>
 F:19-24/Domain: propeptide #status experimental <PRO>
 F:25-607/Product: serum albumin #status experimental <MP>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 4

Query Match

78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2, 1e-154;
 Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAKSEVAFHFKLGEENKALVLAFAQYLQCCPEEDHVKLVNEYTERAKTVADDSAE 60
 |||||
 Db 25 DFKSEAFHFKLGEENKALVLAFAQYLQCCPEEDHVKLVNEYTERAKTVADDSAE 84
 |||||
 QY 61 NCDSLHLEGGDKLVATRETYGEMADCAQOEPRNCFQHHDDNPNLRLRPEV 120
 |||||
 Db 85 GCSKSLHLEGGDKLVATRETYGEMADCAQOEPRNCFQHHDDNPNLRLRPEV 143
 |||||
 QY 121 DVNCTAFHDEETFLKLYEIRARHPYFAPLLEFFAKYKAFTTECCQAADKACLP 180
 |||||
 Db 144 NTLCDERKADKFKWKKYELVETARHPYFAPLLEFFAKYKAFTTECCQAADKACLP 203
 |||||
 QY 181 KLDELDEKASSAKQRLKASQKGERAFKAMAVARSQRPKAFPAFVSKLVYDLTK 240
 |||||
 Db 204 KITEMREKVLASSARQRLKASQKGERAFKAMAVARSQRPKAFPAFVSKLVYDLTK 263
 |||||
 QY 241 VHECCGHDLECCADRADLAKYICENODSISLKECCPKLLEKSHCIAEYENDEMPA 300
 |||||
 Db 264 VHECCGHDLECCADRADLAKYICENODSISLKECCPKLLEKSHCIAEYENDEMPA 323
 |||||
 QY 301 DLPSLADFEYSKDVCKNTAEAKDVEFGMLVEYARHPDYVYLLRLAKYETTLK 360
 |||||
 Db 324 NLPPLADFEYKDVCKNTAEAKDVEFGMLVEYARHPDYVYLLRLAKYETTLK 383
 |||||
 QY 361 CAADPHECYAKYFDEKPLVEEPONLIKONCEFLQGLGYKQNALVRYTKKPVOST 420
 |||||
 Db 384 CAADPHECYAKYFDEKPLVEEPONLIKONCEFLQGLGYKQNALVRYTKKPVOST 443
 |||||
 QY 421 PTLVEYSRNLGKYGKCKCKHPEAKRMPKADYLSVYNOLCVHEPTVSDRYTKCTES 480
 |||||
 Db 444 PTLVEYSRNLGKYGKCKCKHPEAKRMPKADYLSVYNOLCVHEPTVSDRYTKCTES 503
 |||||
 QY 481 LVNRRPFSALVEDETYVPEFNAETFTFHADICTSEKROIKOTALELKHPRKAT 540
 |||||
 Db 504 LVNRRPFSALVEDETYVPEFNAETFTFHADICTSEKROIKOTALELKHPRKAT 563
 |||||
 QY 541 KEOLKAVMDPAFVPEKCKADKKECFKFAEGKLVAAQAL 583
 |||||
 Db 564 EOLKAVMDPAFVPEKCKADKKECFKFAEGKLVAAQAL 606
 |||||

RESULT 6

ABSNS

serum albumin precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999

A:Accession: S06936

R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.

Nucleic Acids Res. 17, 10495, 1989

A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.

A:Reference number: S06936; MUID:90098888; PMID:2602160

A:Accession: S06936

A:Molecule type: mRNA

A:Residues: 1-607 <BRO>

A:Cross-references: EMBL:X17055; NID:g1386; PIDN:CAA3903.1; PID:g1387

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper

teroid hormones (weak bonds with these hormones promote their transfer across the mem

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRO>

F:25-607/Product: serum albumin #status predicted <MAN>

F:29-201/Domain: serum albumin repeat homology <SA1>

F:220-393/Domain: serum albumin repeat homology <SA2>

F:412-591/Domain: serum albumin repeat homology <SA3>

F:27/Binding site: copper (His) #status predicted

F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-39

Query Match

78.4%; Score 2432.5; DB 1; Length 607;

[illegible]

RESULT 11

ABCBS
serum albumin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
R:Hache, R.J.G.; Miskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A>Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CAS>
A:Cross-references: EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748
R:Hache, R.J.G.; Miskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A>Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein submitted to the EMBL Data Library, July 1991
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
A:Cross-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039
R:Rozen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A>Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A:Reference number: A13451; MUID:78019943; PMID:911327
A:Accession: A13451
A:Molecule type: protein
A:Residues: 19-23,'M',25-30 <ROS>
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper mones (weak bonds with these hormones promote their transfer across the membranes), C:Superfamily: Serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:118/Domains: signal sequence #status predicted <SIG>
F:19-26/Product: serum albumin #status predicted <PRO>
F:27-613/Product: serum albumin #status predicted <MAT>
F:32-206/Domains: serum albumin repeat homology <SA1>
F:225-358/Domains: serum albumin repeat homology <SA2>
F:417-596/Domains: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-3

Query Match Best Local Similarity 46.7%; Pred. No. 1,4e+95; Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;

QY 3 HKSEVARFKDLDGENFRKALYLIAFAOYLDOCCPEPDHWIKLYNEVEFEAKTCVADESANEC 62
Db 30 HKSIARIRYNDLKKEETRAKAVAMITFAOYLORCSIEGSKLVKYVDLAOKCVANEDAPEC 89

QY 63 DKSLHTLFGBKLCTVATLRLETYGEMADCCAKOEERNECFQLQHKDNPNLPRLVREYVD 121
Db 90 SKPLPSITILDEICQVKELTRDSYGAMADCCSADPERNECFLSFVSQPDFVQPIQRASD 149

QY 122 VMCTAFHDNETFTLKTLVLEIARRHPPTFYAPBELLFFPKRRYKAATTECCQAADRACLLPR 181
Db 150 VTCEYODNRNSFLGFHFYSVARRHPLRYAPALISFVVDENHALQSCCKSDWGACLDTK 209

QY 182 LDLEIRDGKASSAKORFKASCLOKFGEGRAPAFNAVVALRSOPFAEPAEYSKIATYTDLTKV 241
Db 210 EIYVRKAAKGYSAVQQOIFCGILLKQFGRGVROAROLITLSQYTPAAPSEVSKFPHDSIGV 269

QY 242 HTTECHCDLLECADDRADLAKYTIENDSDISSKILECCCEKPLLEKSHCIAEVENDPAD 301
Db 270 HKECCBEDDMVECMODMARMSNLCSQGDVFSGKITKDCCEPIYVERSGCINAEERDEKPAD 329

QY 302 LPSLIADPVESKDVCNKYAEAKDYFLGMFLYEYARRRPDISVULLRLAKTYETTLEKC 361

A.Residues: 19-45:63-97;102-107;122-184;187-249;255-489;507-609 <TEC>
A.Yachnin, S.; Hsu, R.; Heinrikson, R.L.; Miller, J.B.
Biochim. Biophys. Acta 493, 418-428, 1977
A.Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A.Reference number: A90624; MUID:77242506; PMID:70228
A.Accession: A90624
A.Molecule type: protein
A.Residues: 'S',20-22,'S',24-35 <YAC>
A.Note: dimeric and trimeric forms have been found in addition to the monomeric form
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A.Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A.Reference number: A90757; MUID:78001160; PMID:71198
A.Accession: A90757
A.Molecule type: protein
A.Residues: 'S',20-30,'A',32-37,'A' <AOY>
R.Ruoslahti, E.; Pihko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Kontinen, A.
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
A.Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
A.Reference number: A93042; MUID:75018719; PMID:4138095
A.Accession: A93042
A.Molecule type: protein
R.Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamacki, T.
J. Biol. Chem. 260, 5055-5060, 1985
A.Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A.Reference number: A92520; MUID:85182629; PMID:2580830
A.Contents: annotation; gene, exons and introns
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A.Title: Copper(II)-binding ability of human alpha-fetoprotein.
A.Reference number: A90758; MUID:79001617; PMID:80265
A.Contents: annotation; metal binding
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A.Title: alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A.Reference number: A90759; MUID:80001710; PMID:88900
A.Contents: annotation; bilirubin binding
C.Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma c
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AF
C.Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
properties.
C.Genetics:
A.Gene: GDB:AFP
A.Cross-references: GDB:119660; OMIM:104150
A.Map position: 4q11-4q13
A.Intons: 29/1, 46/2, 90/3, 161/2, 205/3, 238/2, 281/3, 353/2, 397/3, 430/2, 476/3, 551
C:Superfamily: serum albumin; serum albumin repeat homology; metal binding; plasma
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-609/Product: alpha-fetoprotein #status predicted <MAT>
F:29-302/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:22/Binding site: copper (His) #status experimental
F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F:249/Binding site: bilirubin (Lys) #status predicted
F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.3%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.9%; Pred. No. 3.5e-75;
Matches 233; Conservative 116; Mismatches 231; Indels 7; Gaps 3;

3 HKSE-----VAHREFDLGEENKALVLLAFAYLOQCFEDHVKLVNEVEFAKTCVAD 57
DB HNEEGIASILDYSYCTAIEISLADLATIFFAQFVEATYKESKWKAKALTAIEKPTGDE 81
QY SAENCDKSLHTLFGKCLCTVATLRETYGEMADCAKOEPRNECFLOHKNDP-NLPLV 116
DB OSSGCLGNOLPAFLTELCEHEKEILEKYGH-SPDCSGSEGRHNCFLAKKPPASIPLEQ 140
QY RPVEDVMCTAFHNDNEETFLKRYLYEIAARRHPPFYAPPELLFPKRYKAAFTGCCQAADKRA 176

DB VEPVTSCEAYEEDRETFMKNFEIARRHPELVAPJLLMARRYDKIIPSCKAEMAVE 200
QY CLLPKXDELRODGKASSAKORLKASLOKFGERRAKANAVARLSORPFAEFAEVSKLVT 236
DB CQOTAAVYTKRLRESSLLNGACAVMKNFGRTQALITVYKLSOKFTKVNTELOKLYL 260
QY DLTKVTECCGDLLECADRADLAKYICENODSISKLECCERPLEKSHCIAEVEND 296
DB DVAVHEHCRCRDVLDLQDGEKIMSYICSGQDITLNNKITECCKLTLTERGGCIHAEND 320
QY EHPADLPISLADFVSKDYCKRYAKOVFLGMFLYEXARRHDPDSVYLLRLATYERT 356
DB EKEPEGLSPNLNRELFCDROPNDSSEKNEFLASFVHERRRHPLAVSVLITVAAGYDEL 380
QY LEKCCAAADPHCEYAKVDEFEKPLVEEPONLKWCEFEOLGEYKFNALLVRYTKVP 416
DB LEKCPOTENPLECOKGEELQYIOEGALAKRSCGLFQKGEYVYLNQALFVATTKAP 440
QY OVSTPLVEVSNLKGKSCCKHPDAPKMPACEDYLSVYLNQLCVLHEKTPVSDRVTKC 476
DB QLTSELMATFRKMATATATCCOLSEDKILACGEAADIILIGHLCIRHEMPFVNGVQC 500
QY CHESLVNRRPCSALEVEDETVYKPEMNAETPFHADICTLSEKROIKOTALVEYVHK 536
DB CTSSTANRRPCSSLVVDETVYPAFSDCKFLFHKDLQAGQVALQTMKOEFLINDVKOK 560
QY PKATREOLKAVMDPEAFVEKCKADDETCFAEGSKLVASQAALGL 585
DB PQTEDQLEAVIADPSGLEKCCQGEQVCEVCAEFGKIKSTRALGV 609

RESULT 14

PFOO

alpha-fetoprotein precursor - gorilla

C:Species: Gorilla gorilla (gorilla)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999

C:Accession: A37970

R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.

Genomics 9, 60-72, 1991

A>Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primat

A.Reference number: A37970; MUID:91169517; PMID:1706310

A.Accession: A37970

A.Molecule type: DNA

A.Residues: 1-609 <RYA>

A.Cross-references: GB:M8272; MID:9817963; PIDN:AAA73520.1; PID:9177041

C:Genetics:

A.Map position: 4q11-12

A.Intons: 29/1, 46/2, 90/3, 161/2, 205/3, 238/2, 281/3, 353/2, 397/3, 430/2, 476/3;

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-609/Product: alpha-fetoprotein #status predicted <MAT>

F:29-302/Domain: serum albumin repeat homology <SA1>

F:221-394/Domain: serum albumin repeat homology <SA2>

F:413-592/Domain: serum albumin repeat homology <SA3>

F:22/Binding site: copper (His) #status predicted

F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-

F:249/Binding site: bilirubin (Lys) #status predicted

F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 1242.5; DB 1; Length 609;
Best Local Similarity 39.6%; Pred. No. 1e-74;
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;

3 HKSE-----VAHREFDLGEENKALVLLAFAYLOQCFEDHVKLVNEVEFAKTCVAD 57
DB HNEEGIASILDYSYCTAIEISLADLATIFFAQFVEATYKESKWKAKALTAIEKPTGDE 81
QY SAENCDKSLHTLFGKCLCTVATLRETYGEMADCAKOEPRNECFLOHKNDP-NLPLV 116
DB OSSGCLGNOLPAFLTELCEHEKEILEKYGH-SPDCSGSEGRHNCFLAKKPPASIPLEQ 140
QY RPVEDVMCTAFHNDNEETFLKRYLYEIAARRHPPFYAPPELLFPKRYKAAFTGCCQAADKRA 176

```

Db      141 VEPATVSCAEYEDREFENFKFIYELARRHPFLYAPTILLMARHYOKIIPSCCKAENAVE 200
Qy      177 CLLPKDELDEKASSAKORLCAISLOKGEBAFAMAVARLSQRPKAEFEVSKLVY 236
Db      201 CFQTKATVYKELRESLLNQHMCAVMKNKNGTRTFQAIYVTKLSOKFTKYNFTIEIOKLV 260
Qy      237 DLKRVTECCGHDLLLECCADRADLAVYICENODSISKLEKCECEKPLEKSHCIAEVED 296
Db      261 DVANVHEHCRCRGVLDLQDGEKIMTYISQDPLTSKLTTECKKLTTLLEGQCIHAEND 320
Qy      297 EMPADPLSLADEVESKDYCNKAYAEKADVFLGMEFLXYEARHDPYSVLLRLAKTYETT 356
Db      321 EKREGSLPNLNRFLGRDQNFQSSGKKNIFLASFVHEYSRRHQVLAVSLYRAKGVQEL 380
Qy      357 LKCCAAADPHCEYAVFDEFKPLVEBPONLIRKONCELPQDGEYKFNALVRYTKVP 416
Db      381 LEKCPQENPLEQDQGEELQRYIQESQALAKRSGLPQKLEYYLQNFVLAAYTKAP 440
Qy      417 QVSTPLTVEYSRNLGKVKSCCKHPKAPKAPCAEDYLSVYLNQLCVLEHKTPTVSDRYTKC 476
Db      441 QLTSSLMATTRKMAATATCCOLSEKLLACGGAADITIGHLCIRHEMTVPNPGVQC 500
Qy      477 CTESLVNRRPCFSALEVDETYVPERNAETFTFHADICTLSEKERQIKQOTALVELYKHK 536
Db      501 CTSSVANRRPCFSSLVDETYVPARSDDKFTFHKDLCOAQGVALQTMQKQEPFLINLVKOK 560
Qy      537 PKATKQQLKAVMDFAAFVCKCKKADDEKTCFAEBEGKTLVAASQALGL 585
Db      561 PQTTEQLETVIADFSGLEKCCQGOGEVCFABEGOKLISKRTALGV 609

```

RESULT 15

ABXL72

74k albumin precursor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999

C:Accession: B41682; S02693; A05288

R:Koskitts, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.

Mol. Endocrinol. 3, 464-473, 1989

A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.

A:Reference number: A41682; MID:89313788; PMID:2747653

A:Accession: B41682

A:Molecule type: mRNA

A:Residues: 3607 <MOS>

A:Cross-references: GB:M21442; NID:g213930; PIDN:AAA49637.1; PID:g213931

R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.

J. Mol. Biol. 199, 83-93, 1988

A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deletion of the 5'-proximal exon region of the 74-kDa albumin gene results in a nonfunctional gene.

A:Reference number: S02692; MID:88172470; PMID:2451026

A:Accession: S02693

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-48 <SCH>

A:Cross-references: EMBL:226826

R:Molffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.

Eur. J. Biochem. 146, 489-496, 1985

A:Title: Deletion of a transcription of Xenopus 74-kDa albumin genes and destabilization of the 74-kDa albumin gene.

A:Reference number: A05288; MID:85126974; PMID:3971963

A:Accession: A05288

A:Molecule type: mRNA

A:Residues: 459-502, 'L', 504-557 <MOL>

A:Cross-references: GB:M28276

A>Note: The authors translated the codon TAT for residue 63 as Thr

C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thy

C:Genetics:

A:Introns: 27/1

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRO>

F:25-607/Product: 74k serum albumin #status predicted <MAT>
 F:33-201/Domain: serum albumin repeat homology <SAL>
 F:220-393/Domain: serum albumin repeat homology <SAL2>
 F:412-591/Domain: serum albumin repeat homology <SAL3>
 F:30/Binding site: copper (His) #status predicted
 F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

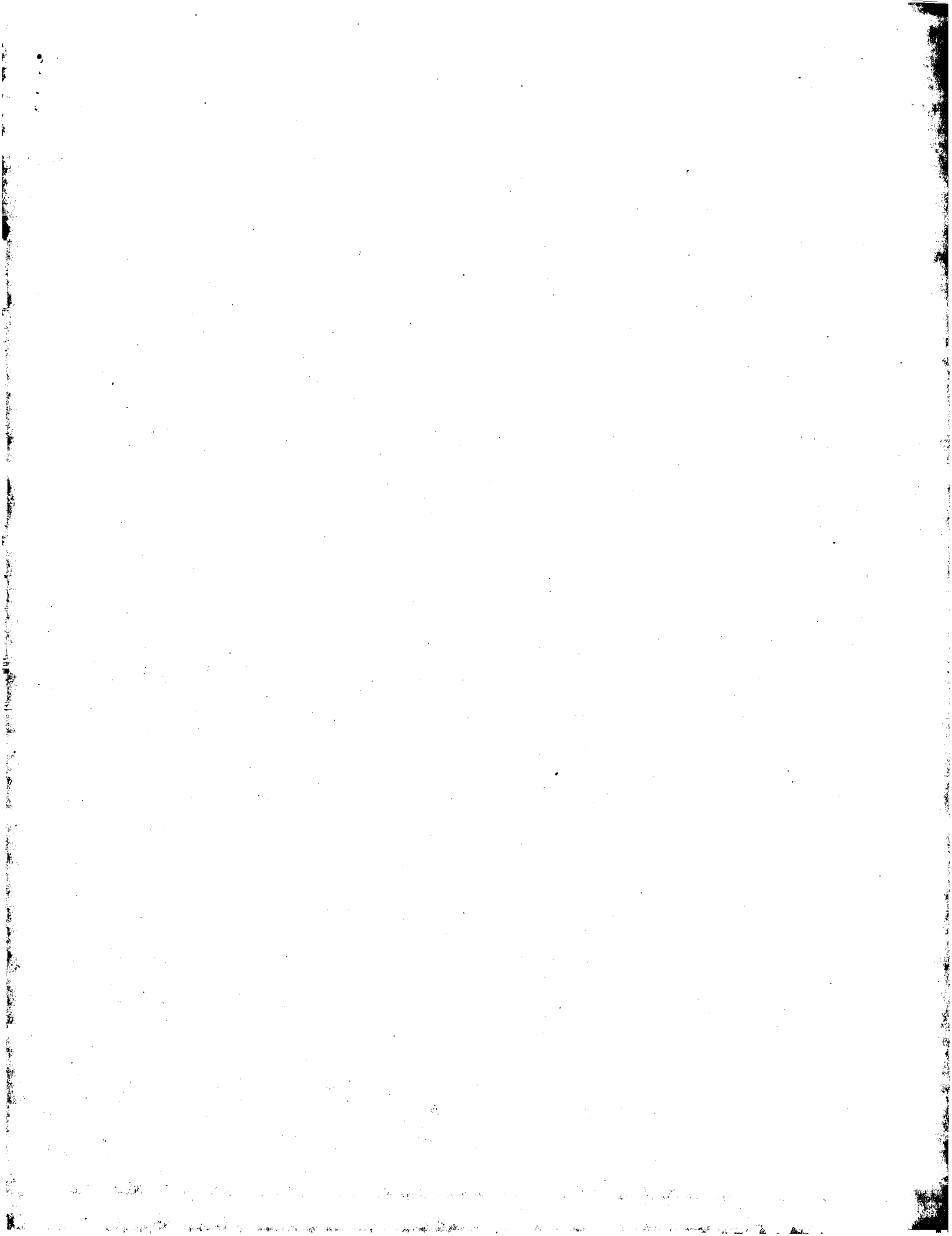
Query Match 38.88; Score 1205; DB 1; Length 607;
 Best Local Similarly 39.38; Pred. No. 3e-72;
 Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

```

Qy      3 HKSEVAFHFKDLDEENFKALVLLAFQYLOQCFEDHVKLVNVEYTERAKTVADESANG 62
Db      30 HHHKADVYIALERFTFGTLTIVSONLQKCSLEELSKLVNEINDPAKSCINDKTPC 88
Qy      63 DKSLLHTLFGDKLTQVATLRETYGBMADCCAKOBERNECPLOKNDPNLRLVREVDV 122
Db      89 EKPVGLTFEDKLCADPAVGAVNEMSKKCCAKQDPERAQCFKARHDEHT---SIKPEPE 145
Qy      123 MCTAFHDNETFLAKYIETARRHPYFAPELLFPKRYKAATTECCQAADKAACLLPKL 182
Db      146 TCKLKEHPDILLSAFIHEPARNHPDLYPPAVLALTKYHKLAHCCCEEDKKECSEKM 205
Qy      183 DELREDEKASSAKORLKCASLOKGEBAFAMAVARLSQRPKAEFEVSKLVTDLTQVH 242
Db      206 KQLMKQSHSIEDKQHNFCWILDNFPEKYLKALMLAKVSHRYPAERKLANHTFEVTHFI 265
Qy      243 TECCHDLEECADRADLAVYICENODSISKLEKCEKPLEKSHCIAEVEDENPADL 302
Db      266 KDCCHDMFCMTERLELFTHTQKHDELSSKLEKCNPLRBTCTIYVLENDVPAEL 325
Qy      303 PSIAADPVESKDYCNKAYAEKADVFLGMEFLXYEARHDPYSVLLRLAKTYETTLKCCA 362
Db      326 SQPTTEFTEDPHVCEKAYENNEVEFLGRYLAVASRKQDELSEDFLOSACEKESLNRKCK 385
Qy      363 AADPHCEYAVFDEFKPLVEBPONLIRKONCELPQDGEYKFNALVRYTKVPYSTPT 422
Db      386 TDNPPECYKAGADRFNMEAKERAYLKONCDIIEHGEYLFENELLIRYTKKMPQSDET 445
Qy      423 LVEVSRNLGKVKSCCKHPKAPKAPCAEDYLSVYLNQLCVLEHKTPTVSDRYTKC 482
Db      446 LIGIAHQMADIGEHCCAVPENQHMPCABGDLTILIGMCRQKTFINNVAACIDSY 505
Qy      483 NRRPCFSALEVDETYVPERNAETFTFHADICTLSEKERQIKQOTALVELYKRPATKE 542
Db      506 GMRSCFTALGPDDVDYPPVTDPTFHDDKICTANKEKHOKFLVLIKVSPLKKN 565
Qy      543 QLKAVMDFAAFVCKCKKADDEKTCFAEBEGKTLVAASQ 580
Db      566 HIDECSAEFLKMYQKCTADEHPCEFTKPVLIHCO 603

```

Search completed: July 22, 2003, 11:45:20
 Job time : 25 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:38:03 ; Search time 15 Seconds

(without alignments)
1617.577 Million cell updates/sec

Title: US-09-833-118-18

Sequence: 1 DAHKSVAHAFKDLGLENFK.....TCFAERCKKLVAASQALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	609	1	P02768 homo sapien
3	2620	84.4	608	1	Q28522 macaca mula
4	2562	82.6	608	1	P49064 felis silve
5	2475.5	79.8	607	1	ALBU_CANEA
6	2450.5	79.0	607	1	ALBU_HORSE
7	2446	78.8	608	1	ALBU_BOVIN
8	2432.5	78.4	607	1	ALBU_RABIT
9	2426	78.2	608	1	ALBU_SHEEP
10	2411.5	77.7	605	1	ALBU_RAT
11	2387	76.9	609	1	ALBU_PIG
12	2378	76.6	608	1	ALBU_MERUN
13	1557.5	50.2	615	1	ALBU_MOUSE
14	1253.5	40.4	609	1	ALBU_CHICK
15	1249.5	40.3	609	1	ALBU_PANTR
16	1242.5	40.0	609	1	FETA_HUMAN
17	1205	38.8	607	1	FETA_GORGO
18	1200	38.7	609	1	ALBU_XENLA
19	1164.5	37.5	606	1	FETA_HORSE
20	1084	34.9	605	1	ALBU_XENLA
21	1067	34.4	611	1	FETA_MOUSE
22	1055	34.0	599	1	FETA_RAT
23	944	30.4	611	1	AFAM_MOUSE
24	928	29.9	608	1	AFAM_RAT
25	747.5	24.1	608	1	ALBU_SALSA
26	742.5	23.9	608	1	ALBU_SALSA
27	699	22.5	382	1	ALBU_RANCA
28	440.5	14.2	1423	1	ALBU_PETMA
29	386	12.4	474	1	VTDB_HUMAN
30	381	12.3	476	1	VTDB_HUMAN
31	378	12.2	476	1	VTDB_RABIT
32	372	12.0	472	1	VTDB_MOUSE
33	133.5	4.3	1391	1	MST2_DROHY

ALIGNMENTS

RESULT 1	ALBU_HUMAN	STANDARD	PRT: 609 AA
AC	P02768; Q13140; Q9UTZ0; Q9UHS3; Q9P117; Q9P157; Q9S574;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	01-APR-1990 (Rel. 14, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Serum albumin precursor.		
GN	ALB.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=6196112; PubMed=3009475;		
RA	Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,		
RA	Beattie W.G., Dugaiczky A.;		
RT	"Molecular structure of the human albumin gene is revealed by		
RT	nucleotide sequence within q11-22 of chromosome 4.";		
RL	J. Biol. Chem. 261:6747-6757(1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=82081882; PubMed=6171778;		
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,		
RA	Najarjan R.C., Seeburg P.H., Wilson K.L.;		
RT	"The sequence of human serum albumin cDNA and its expression in E.		
RT	coli.";		
RL	Nucleic Acids Res. 9:6103-6114(1981).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=82105994; PubMed=6275391;		
RA	Dugaiczky A., Law S.W., Dennison O.E.;		
RT	"Nucleotide sequence and the encoded amino acids of human serum		
RT	albumin mRNA.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver.		
RA	Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;		
RA	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal liver.		
RA	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,		
RT	He F.;		
RT	"Functional prediction of the coding sequences of 79 new genes deduced		
RT	by analysis of cDNA clones from human fetal liver.";		
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.		
RN	[6]		
RP	SEQUENCE OF 25-609.		
RX	MEDLINE=76187907; PubMed=1225573;		
RA	Meloun B., Moravsek V., Kostka V.;		
RT	"Complete amino acid sequence of human serum albumin.";		
RL	FEBS Lett. 58:134-137(1975).		
RN	[7]		

34	132.5	4.3	2230	1	G0G4_HUMAN	013439 homo sapien
35	129.5	4.2	1972	1	MYHB_HUMAN	P35749 homo sapien
36	129	4.2	3210	1	CENE_HUMAN	P49454 homo sapien
37	128	4.1	1005	1	RAS0_MERJA	058718 methanococ
38	126.5	4.1	1972	1	MYHB_RABIT	P35748 oryctolagus
39	126	4.1	1189	1	SCIL_CHICK	090988 gallus gall
40	125	4.0	3038	1	TRIO_HUMAN	075962 homo sapien
41	124.5	4.0	1790	1	USO1_YEAST	P25386 saccharomyc
42	124	4.0	2653	1	CENE_HUMAN	002224 homo sapien
43	123	4.0	1057	1	EGS_HUMAN	P52732 homo sapien
44	123	4.0	1940	1	MYH3_RAT	P12847 rattus norv
45	122.5	3.9	344	1	MST1_DROHY	008695 drosophila

RP SEQUENCE OF 25-609.
 RA Brown J.R., Shockey P., Behrens P.Q.;
 RL (In) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 RL Pergamon Press, New York (1979).
 RN [8]
 RP SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Menaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 RN [9]
 RP PARTIAL SEQUENCE FROM N.A. (PRO2619/PRO2044/PRO1708/PRO2675).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver.";
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RN [10]
 RP SEQUENCE OF 1-26 FROM N.A.
 RA MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 RL regions and the polymorphic gene transcripts.";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [11]
 RP SEQUENCE OF 222-229.
 RA MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 RL acetylsalicylic acid.";
 RL FEBS Lett. 66:173-175(1976).
 RN [12]
 RP SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RA MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RL 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 RN [13]
 RP SEQUENCE OF 409-609 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 RN [14]
 RP DISULFIDE BONDS.
 RA Sabat M.A., Stockbauer P., Moravsek L., Meloun B.;
 RT "Disulfide bonds in human serum albumin.";
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).
 RN [15]
 RP BILIRUBIN-BINDING SITE.
 RA MEDLINE=78186630; PubMed=656055;
 RA Jacobsen C.;
 RT "Lysine residue 240 of human serum albumin is involved in high-
 RT affinity binding of bilirubin.";
 RL Biochem. J. 171:453-459(1978).
 RN [16]
 RP VARIANT CANTERBURY.
 RA MEDLINE=87157744; PubMed=3828358;
 RA Brennan S.O., Herbert P.;
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
 RL domain of serum albumin.";
 RL Biochim. Biophys. Acta 912:191-197(1987).
 RN [17]
 RP VARIANTS NAG-2 AND NAG-3.
 RA MEDLINE=88068523; PubMed=3479777;
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
 RA Satoh C., Neel J.V.;
 RT "Amino acid substitutions in inherited albumin variants from
 RL Amerindian and Japanese populations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
 RN [18]

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
 RA MEDLINE=89345611; PubMed=2762316;
 RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;
 RT "Point substitutions in Japanese allolbumins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
 RN [19]
 RP VARIANTS MANAUS; OSAKA; NAGOTA; FUKUOKA; HONOLULU AND NEW-GUINEA.
 RA MEDLINE=90115905; PubMed=2404284;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RT "Point substitutions in albumin genetic variants from Asia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
 RN [20]
 RP VARIANT REDHILL.
 RA MEDLINE=90115852; PubMed=2104980;
 RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
 RT human serum albumin whose precursor has an aberrant signal peptidase
 RT cleavage site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
 RN [21]
 RP VARIANTS TORINO; VARESE AND VIBO VALENTIA.
 RA MEDLINE=91062352; PubMed=2247440;
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
 RA Watkins S., Putnam F.W.;
 RT "Mutations in genetic variants of human serum albumin found in
 RT Italy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
 RN [22]
 RP VARIANT VENEZIA.
 RA MEDLINE=91296740; PubMed=2068071;
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
 RA Minchiotti L., Putnam F.W.;
 RT "A donor splice mutation and a single-base deletion produce two
 RT carboxyl-terminal variants of human serum albumin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
 RN [23]
 RP VARIANTS IOWA CITY AND KOMAGOME.
 RA MEDLINE=92052189; PubMed=1946412;
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
 RA Matsuda Y., I., Amaki I., Putnam F.W.;
 RT "Genetic variants of serum albumin in Americans and Japanese.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
 RN [24]
 RP VARIANT CASEBROOK.
 RA MEDLINE=91316157; PubMed=1859851;
 RA Peach R.J., Brennan S.O.;
 RT "Structural characterization of a glycoprotein variant of human serum
 RL albumin: albumin Casebrook (494 Asp-->Asn).";
 RL Biochim. Biophys. Acta 1097:49-54(1991).
 RN [25]
 RP VARIANTS SONDRIO AND PARIS-2.
 RA MEDLINE=92190239; PubMed=1347703;
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
 RA Rochu D., Porta F.;
 RT "Two allolbumins with identical electrophoretic mobility are produced
 RT by differently charged amino acid substitutions.";
 RL Biochim. Biophys. Acta 1119:232-238(1992).
 RN [26]
 RP VARIANTS MALMO.
 RA MEDLINE=92390419; PubMed=1518850;
 RA Carlsson J., Sakamoto Y., Laurell C.-B., Madison J., Watkins S.,
 RA Putnam F.W.;
 RT "Allolbuminemia in Sweden: structural study and phenotypic
 RT distribution of nine albumin variants.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8225-8229(1992).
 RN [27]
 RP VARIANT HERBORN.
 RA MEDLINE=93292504; PubMed=8513793;
 RA Minchiotti L., Galliano M., Zapponi M.C., Tenni R.;
 RT "The structural characterization and bilirubin-binding properties of
 RT albumin Herborn, a [Lys240-->Glu] albumin mutant.";
 RL Eur. J. Biochem. 214:437-444(1993).
 RL [18]

Query Match 100.0%; Score 3103; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 2,7e-195;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGSENFKALVLAFAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60
 |||||
 DB 25 DAHKSEVAHREFKDLGSENFKALVLAFAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 84
 |||||
 QY 61 NCDKSLHTLPFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVPEY 120
 |||||
 DB 85 NCDKSLHTLPFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVPEY 144
 |||||
 QY 121 DVMCTAFHNEEFELTKLYEITARRHPYAPPELLFFARKYKAFEECCOADAACACLP 180
 |||||
 DB 145 DVMCTAFHNEEFELTKLYEITARRHPYAPPELLFFARKYKAFEECCOADAACACLP 204
 |||||
 QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSOREPKAFEAESKLVTDLRK 240
 |||||
 DB 205 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSOREPKAFEAESKLVTDLRK 264
 |||||
 QY 241 VHTECCHGDLLECADDRADLAKYICENODSISKLECCCKPLEKSHCIAEVNDMPA 300
 |||||
 DB 265 VHTECCHGDLLECADDRADLAKYICENODSISKLECCCKPLEKSHCIAEVNDMPA 324
 |||||
 QY 301 DLPSLADEVESKDYCKNYAEAKDYVLGFMFLYEYARRHPDYSVLLRLAKTYETLEKC 360
 |||||
 DB 325 DLPSLADEVESKDYCKNYAEAKDYVLGFMFLYEYARRHPDYSVLLRLAKTYETLEKC 384
 |||||
 QY 361 CAAADHECYAKVFEDEKPLVEEPQNLKONCELFOLGEYKFNALLVRYTKRQVOST 420
 |||||
 DB 385 CAAADHECYAKVFEDEKPLVEEPQNLKONCELFOLGEYKFNALLVRYTKRQVOST 444
 |||||
 QY 421 PTLVEVSRLGKVGSKCKHPREKRCADYLSVNLQCVLHETPPSDRYTKCTES 480
 |||||
 DB 445 PTLVEVSRLGKVGSKCKHPREKRCADYLSVNLQCVLHETPPSDRYTKCTES 504
 |||||
 QY 481 LVNRRPCFSALEDETYVPEKFNATFTFHADICTSEKEROIKKOTALVELKHKPKAT 540
 |||||
 DB 505 LVNRRPCFSALEDETYVPEKFNATFTFHADICTSEKEROIKKOTALVELKHKPKAT 564
 |||||
 QY 541 KEOQLKAVMDFAAFVEKCKKADKETCFABEGKRLVAASQALGL 585
 |||||
 DB 565 KEOQLKAVMDFAAFVEKCKKADKETCFABEGKRLVAASQALGL 609
 |||||

RESULT 2
 ALBU_MACMU STANDARD; PRT; 600 AA.

AC Q28532;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 GN ALB.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211971; PubMed=8460152;
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
 RA Dwyer J., Putnam F.W.;
 RT "CDNA and protein sequence of polymorphic macaque albumins that differ
 RT in b1/b2/b3 binding";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
 CC -1- FUNCTION: SERUM ALBUMIN. THE MAIN PROTEIN OF PLASMA. HAS A GOOD
 CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
 CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 DR EMBL: M90463; AAA36906.1; -
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper.
 FT NON_TER 1 1
 FT SIGNAL <1 10
 FT PROPEP 11 16
 FT CHAIN 17 600
 FT DOMAIN 17 197
 FT DOMAIN 204 389
 FT DOMAIN 396 587
 FT METAL 19 19
 FT BINDING 256 256
 FT DISULFID 69 78
 FT DISULFID 91 107
 FT DISULFID 106 117
 FT DISULFID 140 185
 FT DISULFID 184 193
 FT DISULFID 216 262
 FT DISULFID 261 269
 FT DISULFID 281 295
 FT DISULFID 294 305
 FT DISULFID 332 377
 FT DISULFID 376 385
 FT DISULFID 408 454
 FT DISULFID 453 464
 FT DISULFID 477 493
 FT DISULFID 492 503
 FT DISULFID 530 575
 FT DISULFID 574 583
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 94.8%; Score 2942; DB 1; Length 600;
 Best Local Similarity 93.5%; Pred. No. 8.3e-185;
 Matches 345; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSEVAHREFKDLGSENFKALVLAFAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60
 |||||
 DB 17 DFKHSEVAHREFKDLGSENFKALVLAFAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 76
 |||||
 QY 61 NCDKSLHTLPFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVPEY 120
 |||||
 DB 77 NCDKSLHTLPFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVPEY 136
 |||||
 QY 121 DVMCTAFHNEEFELTKLYEITARRHPYAPPELLFFARKYKAFEECCOADAACACLP 180
 |||||
 DB 137 DVMCTAFHNEEFELTKLYEITARRHPYAPPELLFFARKYKAFEECCOADAACACLP 196
 |||||
 QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSOREPKAFEAESKLVTDLRK 240
 |||||
 DB 197 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSOREPKAFEAESKLVTDLRK 256
 |||||
 QY 241 VHTECCHGDLLECADDRADLAKYICENODSISKLECCCKPLEKSHCIAEVNDMPA 300
 |||||
 DB 257 VHTECCHGDLLECADDRADLAKYICENODSISKLECCCKPLEKSHCIAEVNDMPA 316
 |||||

OY 301 DLPSIADVEKSDVCKNYAKADVFLGMLFYEARRRHDYSVLLRLARTYETTTLEKC 360
D 317 DLPISADVEKSDVCKNYAKADVFLGMLFYEARRRHDYSVLLRLARTYETTTLEKC 376
OY 361 CAADPHHCYAKVDEFPPLVEBPONLTKONCELFEOIGEKFFOALLVRYTKKVPQVST 420
D 377 CAADPHHCYAKVDEFPPLVEBPONLTKONCELFEOIGEKFFOALLVRYTKKVPQVST 436
OY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQVLHEKTPVSDRYTKCTES 480
D 437 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQVLHEKTPVSDRYTKCTES 496
OY 481 LVNRRPCSALEVDYETVPKPEFNAETFFHADICTLSKEKQIKQIALVELVKKPKAT 540
D 497 LVNRRPCSALEVDYETVPKPEFNAETFFHADICTLSKEKQIKQIALVELVKKPKAT 556
OY 541 KEOLKAVDDPFAAEVCKCKADDEKTCFAEGSKKIVAAQAAL 583
D 557 KEOLKAVDDPFAAEVCKCKADDEKTCFAEGSKKIVAAQAAL 599

RESULT 3

ALBU_FELCA STANDARD: PRT: 608 AA.

AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID:9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96194824; PubMed-8647469; Hentges F.;
RA Hilger C., Grigioni F., Kohlen M.,
RL "Sequence of the gene encoding cat (Felis domesticus) serum albumin."
Gene 169:295-296(1996).
CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X84842; CA59279.1;
DR HSSP: P02768; IETB.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum_albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 3.
KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
KW Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 BY SIMILARITY.
FT DOMAIN 25 205 SERUM ALBUMIN.
FT DOMAIN 212 397 ALBUMIN 1.
FT DOMAIN 404 595 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.

FT METAL 27 27 COPPER.
FT DISULFD 77 86 BY SIMILARITY.
FT DISULFD 99 115 BY SIMILARITY.
FT DISULFD 114 125 BY SIMILARITY.
FT DISULFD 148 193 BY SIMILARITY.
FT DISULFD 192 201 BY SIMILARITY.
FT DISULFD 224 270 BY SIMILARITY.
FT DISULFD 269 277 BY SIMILARITY.
FT DISULFD 289 303 BY SIMILARITY.
FT DISULFD 302 313 BY SIMILARITY.
FT DISULFD 340 385 BY SIMILARITY.
FT DISULFD 384 393 BY SIMILARITY.
FT DISULFD 416 462 BY SIMILARITY.
FT DISULFD 461 472 BY SIMILARITY.
FT DISULFD 485 501 BY SIMILARITY.
FT DISULFD 500 511 BY SIMILARITY.
FT DISULFD 538 583 BY SIMILARITY.
FT DISULFD 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F605F CRC64;

Query Match 84.4%; Score 2620; DB 1; Length 608;
Best Local Similarity 82.0%; Pred. No: 8.4e-164;
Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

OY 1 DAHSEVAAHRRKDLGEENFKALVLAERAOYLQCCPEEHVYKLVNDETEFAKTCVADESAE 60
D 25 EAHOSEIAHRRNDLGEHFRGLVAVFQYLOCCPEEDHVKLVNDETEFAKTCVADESAA 84
OY 61 NCDSLHLPFGDKICTVATLTREYGMADCCAKOPEENEFLOHKDNPMLPRLVREY 120
D 85 NCESLSHLPFGDKICTVATLTREYGMADCCAKOPEENEFLOHKDNPMLPRLVREY 144
OY 121 DVMTAFHNDNEETFLKLYEIAHRRHYPFVAPPELLFEAKRYKAFTTECCQAADRAACILP 180
D 145 DAMCTAFHNDNEETFLKLYEIAHRRHYPFVAPPELLFEAKRYKAFTTECCQAADRAACILP 204
OY 181 KLDELRECKASSAKQIKKASLOKFGRAKRAVAVLSQFRFAEAEISKVLTDLT 240
D 205 KVDALEKREVLASSAKERLKASLOKFGRAKRAVAVLSQFRFAEAEISKVLTDLT 264
OY 241 VHTCECHGDLLECCADDRADLKYICENDSISSKICEKCEPPLKESHCIAEVENDEMPA 300
D 265 IHKECHGDLLECCADDRADLKYICENDSISSKICEKCEPPLKESHCIAEVENDEMPA 324
OY 301 DLPSIADVEKSDVCKNYAKADVFLGMLFYEARRRHDYSVLLRLARTYETTTLEKC 360
D 325 DLPSIADVEKSDVCKNYAKADVFLGMLFYEARRRHDYSVLLRLARTYETTTLEKC 384
OY 361 CAADPHHCYAKVDEFPPLVEBPONLTKONCELFEOIGEKFFOALLVRYTKKVPQVST 420
D 385 CATDDPFAAEVCKCKADDEKTCFAEGSKKIVAAQAAL 444
OY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQVLHEKTPVSDRYTKCTES 480
D 445 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQVLHEKTPVSDRYTKCTES 504
OY 481 LVNRRPCSALEVDYETVPKPEFNAETFFHADICTLSKEKQIKQIALVELVKKPKAT 540
D 505 LVNRRPCSALEVDYETVPKPEFNAETFFHADICTLSKEKQIKQIALVELVKKPKAT 564
OY 541 KEOLKAVDDPFAAEVCKCKADDEKTCFAEGSKKIVAAQAAL 583
D 565 KEOLKAVDDPFAAEVCKCKADDEKTCFAEGSKKIVAAQAAL 607

RESULT 4

ALBU_CANFA STANDARD: PRT: 608 AA.

AC P49822; 077705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Serum albumin precursor (Allergen Can f 3).

CC	EMBL; X74045; CAA52194.1; -.
DR	PIR; S34053; ABHOS.
DR	HSSP; P02768; IE7B.
DR	InterPro; IPR000264; Serum_albumin.
DR	Pfam; PF00273; transport_prot; 3.
DR	PRINTS; PR00802; SERUMALBUMIN.
DR	ProDom; PD002486; Serum_albumin; 1.
DR	SMART; SM00103; ALBUMIN; 3.
DR	PROSITE; PS00123; ALBUMIN; 3.
KW	Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal; Copper.
FT	SIGNAL 1 18 BY SIMILARITY.
FT	PROPEP 19 24 BY SIMILARITY.
FT	CHAIN 25 607 SERUM ALBUMIN.
FT	DOMAIN 25 204 ALBUMIN 1.
FT	DOMAIN 211 396 ALBUMIN 2.
FT	DOMAIN 403 594 ALBUMIN 3.
FT	METAL 27 27 COPPER (BY SIMILARITY).
FT	DISULFID 77 86
FT	DISULFID 99 115
FT	DISULFID 114 125
FT	DISULFID 147 192
FT	DISULFID 191 200
FT	DISULFID 223 269
FT	DISULFID 268 276
FT	DISULFID 288 302
FT	DISULFID 301 312
FT	DISULFID 339 384

Query Match	79.8%;	Score 2475.5;	DB 1;	Length 607;
Best Local Similarity	76.3%;	Pred. No. 2,2e-154;		
Matches 445;	Conservative 69;	Mismatches 68;	Indels 1;	Gaps 1;
QY 1	DAHKSEVAHREKDIGEENFKALVLIADFAQYIQCCPFEDHAKLVNEVTEFAATCTYADESAE	60		
DB 25	DTHSEIAHRENDIGEEKHFKGLVLAASQYLOQCCPFEDHAKLVNEVTEFAAKCADESAE	84		
QY 61	NCKSLLTTLTGDKICTYATILRETVGMAADCCAKOKEPENECEFLQHKDNPMLPRLVREV	120		
DB 85	NCDSLSHTLFEQDKLCTVATILRATVYGLADCCOKEPENECEFLTHKHDDHMLPKL-KRP	143		
QY 121	DVMCTAFHDNETFLTKYGLIARIHRHPFYAPELLTFEAKRKAAFTTECCOADAARACILP	180		
DB 144	DAQCAALVQEDDKLTKYLIVARHHRPFYFGRPELLFHAEYKADTECCPADDKLACLIP	203		
QY 181	KLDELREBGKASSAKQRLKCAISLQCFGRATKAMAAVAPLRSRPFKAEPAEYSKLYTDLTK	240		
DB 204	KLDELREILTSAKERLCKSSFOFNGERAVKAMSVALRSQKFPKADPAEYSKIYTDLT	263		
QY 241	VHTCECHDILLECADRDADLAKYTCENODSISSKLCECCPELLEKSHCIAEVENDEMPA	300		
DB 264	VHKCECHDILLECADRDADLAKYICEHODSISGKAKACCDPPLQKSHCIAEYKEDDIPS	323		
QY 301	DLPSLAADFEVSKDCKNYAAKADVFLCMFLYEVARRHPDYSVLLILAKYETTLKRC	360		
DB 324	DLPALAADFAEDKEIKETKHKAKADVFLGFLTEYRRRPDYSVLLILAKYETTLKRC	383		
QY 361	CAADPHICVYAKVFDEFRPLVEEPPNLKONCELEFQLEKPKFONALLVRYTKAPQVST	420		
DB 384	CAADPPACYKRVFQDFTPLVEEPPSLVKKNCDDLEFEEVGEYFONALLVRYTKAPQVST	443		
QY 421	PLTVEVSNLGVSKCKCKHPEAKMPCADYLSVNLQCLYHKTYSORVTKCTES	480		
DB 444	PLTVEIGITLGVSRCKLPSESELPCESENHLALALNRLCLVHKTPTVSEKTYCCYDS	503		
QY 441	LVNRPPCSALEVDVTYVPKPEFNATFPFHADICTLSEKEROIKQOTALVELVKKRPAT	540		
DB 504	LAERPPCSALELDDGIVPKPEKATFTFPHADICTLPEDEKIKQOSALAEVLKRPAT	563		
QY 541	KEQLKAVMDDEFAFEVCKCKADDKETCAEESGKLVAASQAL	583		
DB 504	KEQLKTVLGNPSAFYAKCCGRDRKACPAEESPKVLASSQAL	606		
RESULT 6				
ALBU BOVIN STANDARD; PRT: 607 AA.				
ID	ALBU_BOVIN	STANDARD;	PRT:	607 AA.
AC	P02769;	002787;		
DT	21-JUL-1986	(Rel. 01, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Serum albumin precursor (Allergen Bos d 6).			
GN	ALB.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;			
RL	Submitted (Aug-1991) to the EMBL/Genbank/DBD databases.			
RN	[2]			

RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Barry T., Power S., Gannon F.;
 RL Submitted (Jul-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
 RL Submitted (Jul-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-32.
 RX MEDLINE=80024278; PubMed=488109;
 RA McGillyray R.T.A., Chung D.W., Davie E.W.;
 RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin.";
 RL Eur. J. Biochem. 98:477-485(1979).
 RN [5]
 RP SEQUENCE OF 25-424 AND 429-607.
 RA Brown J.R.;
 RT "Structure of bovine serum albumin.";
 RL Fed. Proc. 34:591-591(1975).
 RN [6]
 RP REVISIONS TO 190-195.
 RA Brown J.R.;
 RN Submitted (APR-1975) to the PIR data bank.
 RN [7]
 RP SEQUENCE OF 402-433.
 RX MEDLINE=82023364; PubMed=7283978;
 RA Reed R.G., Putnam F.W., Peters T. Jr.;
 RT "Sequence of residues 400-403 of bovine serum albumin.";
 RL Biochem. J. 191:867-868(1980).
 RN [8]
 RP SEQUENCE OF 19-28.
 RX MEDLINE=77134075; PubMed=843354;
 RA Patterson J.E., Geller D.M.;
 RT "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin.";
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
 RN [9]
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
 RX MEDLINE=91083649; PubMed=2260975;
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
 RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS.";
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
 RN [10]
 RP SEQUENCE OF 25-41.
 RX MEDLINE=88267456; PubMed=3389500;
 RA Hsieh J.C., Lin F.P., Tam M.F.;
 RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";
 RL Anal. Biochem. 170:1-8(1988).
 RN [11]
 RP SEQUENCE OF 437-451.
 RA Vilbois F.;
 RN Submitted (AUG-1998) to the SWISS-PROT data bank.
 RN [12]
 RP DISULFIDE BONDS.
 RA Brown J.R.;
 RT "Structure of serum albumin: disulfide bridges.";
 RL Fed. Proc. 33:1389-1389(1974).
 CC -1- FUNCTION: SERUM ALBUMIN. THE MAIN PROTEIN OF PLASMA, HAS A GOOD BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES, BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFB/ADB FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M73993; AAA51411.1; -
 DR EMBL: X58989; CAA41735.1; -
 DR EMBL: Y17769; CAA76847.1; -
 DR PIR: A38885; ABBOS.
 DR HSSP: P02768; IEB.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma; Albumin; Metal-binding; Lipid-binding; Repeat; Signal;
 KW Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT METAL 27 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 FT CONFLICT 214 214
 FT CONFLICT 302 302
 FT CONFLICT 304 305
 FT CONFLICT 324 324
 FT CONFLICT 394 395
 FT CONFLICT 437 437
 FT CONFLICT 493 494
 SQ SEQUENCE 607 AA; 69293 MW; 39167DDE76858D4 CRC64;
 A -> T (IN REF. 2 AND 5).
 C -> K (IN REF. 5).
 K P -> P C (IN REF. 5).
 N -> D (IN REF. 5).
 ST -> TS (IN REF. 5).
 K -> R (IN REF. 11).
 SE -> ES (IN REF. 5).
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;
 Best Local Similarity 75.8%; Pred. No. 9,4e-153;
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;
 QY 1 DAHKSVAHFRPDLDGENRKALVLAFAQYLDQCPREDHVKLVNTEYTERAKTCVADSEAE 60
 DB 25 DTHKSEIARHFRPDLDGEHFKGLVLAISOYLDQCPDEHVKLVNTEYTERAKTCVADSEAH 84
 QY 61 NCDKSLHTLFGDGLCVATLRRTYEGMADCCAKOEERNECFLOHDDNNPRLVPRPV 120
 DB 85 GCKSLHTLFGDGLCVASLRTYEGMADCCAKOEERNECFLOHDDNNPRLVPRPV 143
 QY 121 DVNCTAHNDEETFLKLYEIAIRHPYFAPPELLFARRYKNAFTCCQADAKACLP 180
 DB 144 NTLCDKFAKDEKFKWYLYEIAIRHPYFAPPELLYANKYNGVFCQCAEDKAGCLLP 203
 QY 181 KIDELDEKQKASSAKRLCASIQKGEFAFRMANARISQRPKAEFAVSTLYDTRK 240
 DB 204 KIEEMREKVALRCARQRCASIQKGEFAFRMANARISQRPKAEFAVSTLYDTRK 263
 QY 241 VTECHGDLLEACADRADLAKYICENQSISSKLECCERPLLEKSHCIAEVENDMPA 300

```

Db 264 VHKECCHDILLECADRADLAKKYICDNDJTSSKLECCDPRLEKSHCIAEVENDEAIPE 323
Oy 301 DLPSTADAFVESKDVCKNYAEAKDVFLGMFLYEYRRHPDVSVLLRLAKTYENTLEK 360
Db 324 NLPLTADFAEDKDYCKNYOEAKDAFLSGFLYEYRRHPDVSVLLRLAKTYENTLEK 383
Oy 361 CAADPHBCYAKVDEFPRLVEEPONLKONCELEFQEGEYKFNALLVRYTKKVPQVST 420
Db 384 CAKDDPHACYSYVFEKDLKHLVDEPQNLKONCDQEGEYKFNALLVRYTKKVPQVST 443
Oy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTVSDRYTKCSES 480
Db 444 PTLVEVSRLGKVGSKCKCKHPEAKRMPCTEDYLSLILNRLCVLHKEKTVSEKVTCKSES 503
Oy 481 LVNRRPCFSALVEDETVYVPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPRAT 540
Db 504 LVNRRPCFSALVEDETVYVPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPRAT 563
Oy 541 KEQLKAVMDFAAFVEKCKKADDEKTCFAEBSKRLVAASQAL 583
Db 564 EEOQKTVMEVFVAFYDKCCAADDKKACFAVEGPKLVSTQIAL 606

```

RESULT 7

```

ALBU_RABIT STANDARD: PRT: 608 AA.
ID ALBU_RABIT STANDARD: PRT: 608 AA.
AC P49065:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN New Zealand white; TISSUE-Liver;
RC Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: BELONGS TO THE ALB/AFIP/VDB FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U18344; AAB58347.1; -.
DR HSSP: P02768; 1E7B.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum_albumin; 1.
DR SMART: SM00103; ALBUMIN_3.
DR PROSITE: PS00212; ALBUMIN_3.
DR Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; signal;
KW Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.

```

```

FT DOMAIN 212 397 ALBUMIN 2.
FT FT 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68914 MW; CE5E92647AAFE9A2 CRC64;

```

Query Match 78.8%; Score 2446; DB 1; Length 608;
 Best Local Similarity 74.1%; Pred. No. 1.9e-152;
 Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 0;

```

Oy 1 DAHKESEVARRPEKDLCEENFKALVLAFAOYLOOCPEDEHVKLVNEVTEFAKTCVADSEAE 60
Db 25 EAHKSEIAHRRNDVGEHFIQVLTTFESQYLOKCPYEBHAKLVKCVTDLAACAVIDESAA 84
Oy 61 NCDKSLHTLFGDKICTVATLTETYGEMADCCAKOEPEBNECFLOHKDDNPILRLVPEV 120
Db 85 NCDKSLHTLFGDKICTVATLTETYGEMADCCAKOEPEBNECFLOHKDDNPILRLVPEV 144
Oy 121 DVMCTAFNDNETPLAKTLVEIARHRYFAVAPELLFFAKRKKAATTECCQAADKACLLP 180
Db 145 DVLCKAFHDDKAFEGHYLVEARRHPYFAELLVYAKKAYKALLTECECAADKACLLP 204
Oy 181 KLDELRLDEGKASASAKORLCKASLOKFGERRAFKAMAVARLSORFPAEFAEYSKLVTDLT 240
Db 205 KLDALEEGSLISMAOERLRKASIOKFGDRATKAMALVLSQRFPAADTDTSKITYDILT 264
Oy 241 VHTCECHDILLECADRADLAKKYICENODSISKLECCDPRLEKSHCIAEVENDEMPA 300
Db 265 VHKECCHDILLECADRADLAKKYICENODSISKLECCDPRLEKSHCIAEVENDEMPA 324
Oy 301 DLPSTADAFVESKDVCKNYAEAKDVFLGMFLYEYRRHPDVSVLLRLAKTYENTLEK 360
Db 325 GLPVAABEEFVEDKDYCKNYAEAKDVFLGMFLYEYRRHPDVSVLLRLAKTYENTLEK 384
Oy 361 CAADPHBCYAKVDEFPRLVEEPONLKONCELEFQEGEYKFNALLVRYTKKVPQVST 420
Db 385 CATDDPHACYSYVFEKDLKHLVDEPQNLKONCDQEGEYKFNALLVRYTKKVPQVST 444
Oy 421 PTLVEVSRLGKVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTVSDRYTKCSES 480
Db 445 PTLVEVSRLGKVGSKCKCKHPEAKRMPCEVEDYLSVNLQCVLHKEKTVSEKVTCKSES 504
Oy 481 LVNRRPCFSALVEDETVYVPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPRAT 540
Db 505 LVNRRPCFSALVEDETVYVPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPRAT 564
Oy 541 KEQLKAVMDFAAFVEKCKKADDEKTCFAEBSKRLVAASQALG 584
Db 565 NDQLKTVGGEFTALLDKCCSAEDKACFAVEGPKLVSTQIALG 608

```

RESULT 8

```

ALBU_SHEEP STANDARD: PRT: 607 AA.
ID ALBU_SHEEP STANDARD: PRT: 607 AA.
AC P14639:
DT 01-APR-1990 (Rel. 14, Last sequence update)

```

DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 ON NCBI_TaxID=9940;
 RN [1]
 RP SOURCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90098888; PubMed=2602160;
 RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
 RT "Nucleotide and deduced amino acid sequence of sheep serum albumin."
 RL Nucleic Acids Res. 17:10495-10495(1989).
 CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VDL FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17055; CAA34903.1; -.
 DR PIR: S06936; ABSH.
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot. 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin. 1.
 DR SMART: SM00103; ALBUMIN. 3.
 DR PROSITE: PS00212; ALBUMIN. 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 607 SERUM ALBUMIN.
 FT DOMAIN 25 204 ALBUMIN 1.
 FT DOMAIN 211 396 ALBUMIN 2.
 FT DOMAIN 403 594 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 147 192 BY SIMILARITY.
 FT DISULFID 191 200 BY SIMILARITY.
 FT DISULFID 223 269 BY SIMILARITY.
 FT DISULFID 268 276 BY SIMILARITY.
 FT DISULFID 288 302 BY SIMILARITY.
 FT DISULFID 301 312 BY SIMILARITY.
 FT DISULFID 339 384 BY SIMILARITY.
 FT DISULFID 383 392 BY SIMILARITY.
 FT DISULFID 415 461 BY SIMILARITY.
 FT DISULFID 460 471 BY SIMILARITY.
 FT DISULFID 484 500 BY SIMILARITY.
 FT DISULFID 510 517 BY SIMILARITY.
 FT DISULFID 537 582 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 78.4%; Score 2432.5; DB 1; Length 607;
 Best Local Similarity 75.0%; Pred. No. 1,4e-151;
 Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAAHRRKDLGENEFKALVLIAPAOYLQOCPFEDHVKLVNEYTEFAKTGVADSEAE 60
 DB 25 DTHKSELAHRRFNDLGEENRFGVLVIAFSQYLQOCPFEDHVKLVNEYTEFAKTGVADSEAE 84
 QY 61 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAOEPRNECFLOHKDNDNPLRYAPEV 120
 DB 85 GCDKSLHTLFGDELCKVATLRETYGEMADCCOEPRNECFLNHKKDSDPLKLT-KPEP 143
 QY 121 DVMCTAFHNEEFLLKYLLEYIARRHRYFAPLLEFPFAKRYKAFFCECAAKACLLP 180
 DB 144 DFLCAEFKADKKEKFWGKYLLEYIARRHRYFAPLLEFYANKNGVFQCECAEKACLLP 203
 QY 181 KLDELDEGKASAKORLKCASLOKGEERFKMAVARISQRPKAEFAVSKLVYDITK 240
 DB 204 KIDAMREKVLASSARQRLKASLOKGERLAKMSVARLSQKPKADFTVYTIIVDITK 263
 QY 241 VIRECHGDLLECCADRADLAKYICENODSISSEKLECCERPELLESHCIAEYNDMPA 300
 DB 264 VIRECHGDLLECCADRADLAKYICDHQDASSKLECCDQPVLEKSHCIAEYNDMPA 323
 QY 301 DLPSLADPFVESKDVCKNYAEADVPLGMFLYEARHPDYSVLLRLAKYETITLKC 360
 DB 324 NLPEPLADFADEKVECKNYQEAADVPLGSLYESSRHPDYAVSVLLRLAKYETITLKC 383
 QY 361 CAADAPHECYAKYFDEKFPLEPPEPONLIKONCELEFEOLOGEYKPPONALLVRYTKPQVST 420
 DB 384 CAEDHPACATYFDDKILHLYDEPONLIKNCLEFEGHGYGONALIVYTKAPQVST 443
 QY 421 PTLVEYSRNIGKYGKCKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 444 PTLVEISRSIGKYGKCKCAKPESEBMPCTEDYLSLILNRLCVLHEKTPVSEKTYKCTES 503
 QY 481 LVNRRPFSLVDEYETVPRKEFNAETFTFADICTLSEKROIKKQATVLYKHKPKAT 540
 DB 504 LVNRRPFSDTLTDEYETVPRKEFETFHADICTLPDTEKQIKQATVLYKHKPKAT 563
 QY 541 KEOLKAVMDQFAFVEKCKADKCEFAEKGKLVAAQOAL 583
 DB 564 DEOLKYMENFVAFVDCCAADKCEGCVLEGRVLAQOAL 606

RESULT 9
 ALBU RAT STANDARD; PRT; 608 AA.
 ID ALBU RAT
 AC P02770; P11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
 GN ALB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SOURCE FROM N.A.
 RX MEDLINE=8123722; PubMed=7017712;
 RA Sargent T.D., Yang M., Bonner J.;
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA."
 RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
 RN [2]
 RP SEQUENCE OF 1-38, AND PROCESSING.
 RX MEDLINE=77249657; PubMed=893447;
 RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
 piece. Analysis of the direct translation product of albumin
 messenger RNA."
 RL J. Biol. Chem. 252:6846-6855(1977).
 RN [3]
 RP SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen

[illegible]

```

RL Nucleic Acids Res. 16:9045-9045(1988).
CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VPB FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12422; CAA30970.1; -.
DR EMBL; M36787; AAA30988.1; -.
DR PIR; S01382; ABPG.
DR HSSP; P02768; 1E7H.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Transport_prot; 3.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR KMW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
KW COPPER.
FT NON_TER 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 209 202 ALBUMIN 1.
FT DOMAIN 401 592 ALBUMIN 2.
FT METAL 31 31 ALBUMIN 3.
FT DISULFID 75 84 COPPER (BY SIMILARITY).
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1FEE CRC64;
Query Match 77.7%; Score 2411.5; DB 1; Length 605;
Best Local Similarity 76.0%; Pred. No. 3.3e-150;
Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

```

```

QY 181 KLDELDEGKASSAKORLKASLOKFGERRAFKAMAVARLSORPEKAEFAVSKLTVDLR 240
DB 202 KIHLEKELVITSAKORLKASLOKGERAFKAMVARLSORPKADFTFISKIVDLAK 261
QY 241 VHECHGDLLECCADRADLAKYICENODISSKLEKCEKPLLEKSHCIAEYNDMPA 300
DB 262 VHECHGDLLECCADRADLAKYICENODISSKLEKCEKPLLEKSHCIAEYNDMPA 321
QY 301 DLPSIADAPFESNDCKNVAEADVPLGMFLYETARRHPDYSVLLRLAKYETLLEK 360
DB 322 DLNPLEHDEVEDKEVCKNKEADVPLGFLYETARRHPDYSVLLRLAKYETLLEK 381
QY 361 CAADPHCEYAKFDEFKPLPEPOMLIKONCELEFOLGEYKONALVYTKKPOVST 420
DB 382 CAEDDPACATYFDKFPPLVDEPKLIKONCELEFOLGEYKONALVYTKKPOVST 441
QY 421 PLVEVSRNLGKYSKCKHPEAKRMPCADYLSVYLNQCLVLEKTPVSDRYTKCCTES 480
DB 442 PLVEVSRNLGKYSKCKHPEAKRMPCADYLSVYLNQCLVLEKTPVSDRYTKCCTES 501
QY 481 LVNRRPCFSALVEDETVYKPEFNAETFTFADICTSEKROIKKOTALVELYKHPKAT 540
DB 502 LVNRRPCFSALVEDETVYKPEFNAETFTFADICTSEKROIKKOTALVELYKHPKAT 561
QY 541 KEOLKAVMDPFAFVEKCKKADDEKTCFAEGEKKV 576
DB 562 EEOELRVYLGFAFVOKCCAPADPHCEYAKFVEGKPV 597

RESULT 11
ID ALBU_MERUN STANDARD; PRT; 609 AA.
AC Q35090;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_Taxid=10047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=9811663; PubMed=9455485;
RA Yoshida K., Seto-Ohshima A., Shinohara H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RL DNA Res. 4:351-354(1997).
CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VPB FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB006197; BAA21765.1; -.
DR HSSP; P02768; 1E7H.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; Transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.

```



```

Query Match      76.6%; Score 2378; DB 1; Length 608;
Best Local Similarity 72.4%; Pred. No. 5e-148;
Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;

QY      1 DAHRSVAHPRKDEGEENFKALVLAFAQYUQOCPEPDHVKLVNEDTEFAKTCVADESAAE 60
      : ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db      25 EAHKSEIAHRRNDGEQHFKGLVLIASFQYLQKCSYDHNAKLVQAEVDVDFATTCVADESAA 84

QY      61 NCDSKSLHTLFQDGLCTVATLRETYGEMADCCAKOEPRNEECFLHKDKDNPMLPRIVPREV 120
      : ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db      85 NCDSKSLHTLFQDGLCTVATLRETYGEMADCCAKOEPRNEECFLHKDKDNPMLPRIVPREV 144

QY      121 DVMCTAFHDNEETFLKKYVLEIARRHPFYFAPPELLFFFAKRYKAAFTTECCQAADRAACLLP 180
      : ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db      145 EAMCTSFRENFTTGMGHLHEVARHHPFYFAPPELLFYAEGYNELITQCCALADKESCLTP 204

```

OY	181	LLDLEDPBGKSSAKSKORIKKASLQKPGFRAKAAVAARLSQRFPAKAEVSKLVTPLTK	240
DB	205	KLDDVKEKALVSVYRORRKCSSMOKFGRATKAAVAARLSQTFPNADFAETIKLATDITK	264
OY	241	VHTSCCHDILLECCADDPRADLAKYICENODSISSKLKECCCEKPLLEKSHCIAVENDEMPA	300
DB	265	VNKECCCHDILLECCADDPRADLAKYICENODSISSKLQTCCKOKPLKKHCLSEVHDMPA	324
OY	301	DLPLSLADFEVSKDVCNKYAEAKDVLGMLFLEYARRHPDYSVLLRLAKTYETTLK	360
DB	325	DLPLSLADFEVDQCVCKNYAEAKDVLGTFLEYESRRHPDYSVLLRLAKKYEATLEKC	384
OY	361	CAADPHCYAKYVDEKREPLVEEPPONLKKONCELPEDLOGEKKPQNALLRYTKVPOYST	420
DB	385	CAEANPPACXYTVALEPPLVEEPPKNLTKTQCDIEKLGEGFNALLRYTORAPYST	444
OY	421	PTLVEVSNLKGVSKCCCKHEAKRMPCAEDYLSVLNQLCVLHEKTPVSPRVRKCTES	480
DB	445	PTLVEVARNLRGVRGKCTLPEDORLPCVEYELISALINRVCLLEKTPVSHYTKCCSGS	504
OY	541	KEQLKAWMDPFAAEVEKCKKADDEKTCFAEGEKKLVAAQAL	583
DB	565	AEQLKTVMDPFAQELDTCCKAADKDTCESTEGPNLVYTRCKDAL	607

RESULT 13

ALBU_CHICK	STANDARD:	PRT:	615 AA.
ID	ALBU_CHICK		
AC	P19121:		
DT	01-NOV-1990 (Rel. 16, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Serum albumin precursor.		
GN	ALB.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver:		
RA	Cassady A.I., Saliklid C.K., Baverstock P., Wallace J.C.,		
RL	Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE OF 1-28 FROM N.A.		
RX	MEDLINE=83161037; PubMed=6187737;		
RA	Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.,		
RT	"The 5' noncoding and flanking regions of the avian very low density		
RT	apolipoprotein II and serum albumin genes. Homologies with the egg		
RT	white protein genes.";		
RL	J. Biol. Chem. 258:4556-4564(1983).		
RN	[3]		
RP	SEQUENCE OF 19-30.		
RX	MEDLINE=78019943; PubMed=911327;		
RA	Rosen A.W., Geller D.M.,		
RT	"Chicken microsomal albumin: amino terminal sequence of chicken		
RT	prealbumin.";		
RL	Biochem. Biophys. Res. Commun. 78:1060-1066(1977).		
CC	-1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD		
CC	BINDING CAPACITY FOR WATER, CA++ , NA+ , K+ , FATTY ACIDS, HORMONES,		
CC	BIURIBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE		
CC	COLLOIDAL OSMOTIC PRESSURE OF BLOOD.		
CC	-1- SUBCELLULAR LOCATION: Secreted.		
CC	-1- TISSUE SPECIFICITY: PLASMA.		
CC	-1- SIMILARITY: BELONGS TO THE ALB/AFIP/VDL FAMILY.		
CC	-1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 22 22 COPPER AND NICKEL (BY SIMILARITY).
 FT DISULFID 99 114 BY SIMILARITY.
 FT DISULFID 113 124 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 609 AA; 68741 MW; C032987CAD0E672B CRC64;

Query Match 40.4%; Score 1253.5; DB 1; Length 609;
 Best Local Similarity 40.1%; Pred. No. 1,1e-74;
 Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

OY 3 HKSE-----VAHREKDLGSEENFKALVLIAPAOYLOCCPEEDHVKLVNEVEFAKTVADE 57
 DB 22 HRNRYGASLIDSTYQCAETNLVDLATIFRQFOEATYKVEVSMVADALTAIKPGDE 81
 OY 58 SAENCDLSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFGLHKDNDP-NLRLV 116
 DB 82 QASGCLLENOLPAFLIEELCKREKILEKYGH-SDCCOSEEGHNFGLAHKKTPTASIFPQ 140
 OY 117 RPEVDVCTAFHNDNEFFLKKYELIARRHRYAPPELLFPAKYKAFTGCCADAKAA 176
 DB 141 VPEPVTSCEAYEEDRETFMNFKEIYIARRHFLYAPILLMAARYDKIIPSCAKENAVE 200
 OY 177 CLLPKLDLDEGKASSAKORLKCASLOKGEERAFKMAVARSOREPKAFKFAVSKLYT 236
 DB 201 CFQTKATVYKELRESSLNOMHACAVKMGESTRFQALITYKLSOKFTKVFTEIQKLV 260
 OY 237 DLKVTHECHGDLLECCADRADLAKYICENODSISSEKCEKPLEKSHCIAEYEND 296
 DB 261 DVAVHHCRCRGVLDLCLDQEKIMSYICSGQDILSKITPECCCLTTLERQCCIIHEND 320
 OY 297 EMPADPLSLADYVESDVCKNTVAEAKDVLGMLFYEARHDPYSVLLIRLAKTYETT 356
 DB 321 EKPEGLSPNLNRFGLDGFNOFSSEKNIIFLASFVHESRRHPOLAVSVILIRVAKGYOL 380
 OY 357 LEKCAADPHCEYAKYFDEFKPLVERPONLKNCELEFQJGKXKONMLVAYTKVP 416
 DB 381 LEKCFOTENLECCODKKEEELQIKQISQALAKRSCGFQJLGLGYQNAFLVAYTKAP 440
 OY 417 QVSTPFLVEVSRNLKGVSKCKHPEAKRMPKACADYLSVNLQCLVLEKTPVDRYTK 476
 DB 441 QLTSELMATRKMAATAATCCOLSEDKLACGEGADIIIGHCINHETTPVPDVGQC 500
 OY 477 CTSELVNRRCFSALYEDETVYKPEFNAETFTPHADICTLSEKROIKQATLVELYKH 536
 DB 501 CTSSYARRRCFSSLYVDETVYPAFSDCKFIHFKDKLOAGVALQTMKQEFLLINLVOK 560
 OY 537 PKATKEDLKVMDFAFVEKCKKADKKECFEAEKGIKVAASOAALGL 585
 DB 561 PQTTEOLEAVIADFSGLERKCCGQGOEVCFAEEGKLLIKSTRALGV 609

RESULT 15
 FETA_HUMAN
 ID FETA_HUMAN STANDARD: PRT: 609 AA.
 AC P02771;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
 DE AFP.
 GN AFP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-83273664; PubMed-6192439;
 RA Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
 RT "Primary structures of human alpha-fetoprotein and its mRNA."; Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87185438; PubMed-2436661;
 RA Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;
 RT "Structure, polymorphism, and novel repeated DNA elements revealed by a complete sequence of the human alpha-fetoprotein gene."; Biochemistry 26:1332-1343(1987).
 RN [3]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE-93278385; PubMed-7684942;
 RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M., Tilgham S., Krumlauf R., Tuddenham E.G.D.;
 RT "A G-->A substitution in an HNF 1 binding site in the human alpha-fetoprotein gene is associated with hereditary persistence of alpha-fetoprotein (HPAFP)."; Hum. Mol. Genet. 2:379-379(1993).
 RN [4]
 RP SEQUENCE OF 429-556 FROM N.A.
 RX MEDLINE-83158778; PubMed-6187626;
 RA Beattie W.G., Dugaiczky A.;
 RT "Structure and evolution of human alpha-fetoprotein deduced from partial sequence of cloned cDNA."; Gene 20:415-422(1982).
 RN [5]
 RP PARTIAL SEQUENCE OF 19-609
 RX MEDLINE-91242409; PubMed-1709810;
 RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F., Ceccarini C., Terrana B.;
 RT "Human alpha-fetoprotein primary structure: a mass spectrometric study."; Biochemistry 30:5061-5066(1991).
 RN [6]
 RP PRELIMINARY SEQUENCE OF 19-35.
 RX MEDLINE-77242506; PubMed-70228;
 RA Yachnin S., Hsu R., Heinrichson R.L., Miller J.B.;
 RT "Studies on human alpha-fetoprotein. Isolation and characterization of monomeric and polymeric forms and amino-terminal sequence analysis."; Biochim. Biophys. Acta 493:418-428(1977).
 RN [7]
 RP PRELIMINARY SEQUENCE OF 19-38.
 RX MEDLINE-78001760; PubMed-71198;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Comparative chemical structures of human alpha-fetoproteins from fetal serum and from ascites fluid of a patient with hepatoma."; Cancer Res. 37:3663-3667(1977).
 RN [8]
 RP PRELIMINARY SEQUENCE OF 19-39.
 RX MEDLINE-75018719; PubMed-4138695;
 RA Kuoslahti E., Plinko H., Vaheri A., Seppala M., Virolainen M., Kontinen A.;
 RT "Alpha fetoprotein: structure and expression in man and inbred mouse strains under normal conditions and liver injury."; J. Johns Hopkins Med. J. Suppl. 3:249-255(1974).
 RN [9]
 RP GENE STRUCTURE.
 RX MEDLINE-85182629; PubMed-2580830;
 RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G., Tamaoki T.;

RT "The human alpha-fetoprotein gene. Sequence organization and the 5' flanking region." J. Biol. Chem. 260:5055-5060(1985).

RL [10]

RN METAL-BINDING.

RP MEDLINE=79001617; PubMed=80265;

RA Aoyagi Y., Ikenaka T., Ichida F.;

RT "Copper(II)-binding ability of human alpha-fetoprotein.";

RL Cancer Res. 38:3483-3486(1978).

RN [11]

RP BILIRUBIN-BINDING.

RA MEDLINE=80001710; PubMed=89900;

RT Aoyagi Y., Ikenaka T., Ichida F.;

RT "Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability.";

RL Cancer Res. 39:3571-3574(1979).

RN [12]

RP SULFATION.

RA MEDLINE=86042625; PubMed=2414772;

RT Liu M.C., Yu S., Sy J., Redman G.M., Lipmann F.;

RT "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";

RU Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).

CC -1- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.

CC -1- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.

CC -1- TO THE MONOMERIC FORM.

CC -1- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.

CC -1- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD. REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.

CC -1- PTA: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE NATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.

CC -1- PTM: SULFATED.

CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.

CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sdb.ch/announce/](http://www.isb-sdb.ch/announce/or send an email to license@sdb-sdb.ch) or send an email to license@sdb-sdb.ch).

CC -----

DR EMBL: M10949; AAA51674.1; -

DR EMBL: M10950; AAA51675.1; -

DR EMBL: V01514; CAA24758.1; -

DR EMBL: M16110; AAB58754.1; -

DR EMBL: Z19532; CAA79592.1; -

DR PIR: A03234; FPHO.

DR PIR: A26624; A26624.

DR HSSP: P02768; 1E7B.

DR GLYCOSULEDB: P02771; -

DR Stena-2DPAGE; P02771; -

DR GeneW; HGNC:317; AFP.

DR MIM: 104150; -

DR InterPro: IPR000264; Serum_albumin.

DR Pfam: PF00273; transport_prot; 3.

DR PRINTS: PR00802; SERUMALBUMIN.

DR PRODOM: PD002486; Serum_albumin; 1.

DR SMART: SM00103; ALBUMIN; 3.

DR PROSITE: PS00212; ALBUMIN; 2.

DR Glycoprotein; Sulfation; Albumin; Plasma; Repeat; Metal-binding;

KA Copper; Nickel; Signal; Polymorphism.

FT SIGNAL 1 18

FT CHAIN 19 609 ALPHA-FETOPROTEIN.

FT	DOMAIN	20	205	ALBUMIN 1.
FT	DOMAIN	212	397	ALBUMIN 2.
FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	22	22	COPPER AND NICKEL.
FT	DISULFID	99	114	
FT	DISULFID	113	124	
FT	DISULFID	148	193	
FT	DISULFID	192	201	
FT	DISULFID	224	270	
FT	DISULFID	269	277	
FT	DISULFID	289	303	
FT	DISULFID	302	313	
FT	DISULFID	384	393	
FT	DISULFID	416	462	
FT	DISULFID	461	472	
FT	DISULFID	485	501	
FT	DISULFID	500	511	
FT	DISULFID	538	583	
FT	DISULFID	582	591	
FT	CARBOHYD	251	251	
FT	VARIANT	570	570	
FT	SEQUENCE	609 AA: 68677 MM: 4D4E45820E1C2D4F CRC64:		
FT	Query Match	40.3%; Score 1249.5; DB 1; Length 609;		
FT	Best Local Similarity	39.9%; Pred. No. 1.9e-74;		
FT	Matches	235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;		
QY	3 HKSE-----VAHREFDGEENFKALVLAFAQYLOOCPEHNVKLVNEYTEFAKTVAD 57			
DB	22 HNEGYIASILDSYCTAEISLADLTTFEAFVGEATYKSKVKALPAIEKPTGDE 81			
QY	58 SAENCDKSIHLTFGDKTAVATLRETYGEMADCCAKOBERNECFLOKHDND-NLPLRV 116			
DB	82 OSSGCLNOLPAFLFEELCHEKILEKIGH-SPCCSSEBGRNCLAKKPPASIPLEQ 140			
QY	117 PREVVMCTAFHNDNEFTLKKYIEIARRHPYFAPELLFFAKRYKAAFECCQADRAA 176			
DB	141 VEPYTSCEAVYEDRETFMKNFIETARRHPYFAPELLFFAKRYKAAFECCQADRAA 200			
QY	177 CLPLRYDELDRGKSSAKQRIKASLOKFGRRAKAAVARSFRKAEFAEVSQIVT 236			
DB	201 CQTAATATYVELRSSLINOHACAVMNFGRTOALTIVKLSOKFTKVNTEIOKLV 260			
QY	237 DLTQHTCCGHDLLECADRADLAKYICENODSISKLECCERPLEKSHCIAEVND 296			
DB	261 DVANHEHRCRDYLDLQDQGEKIMSYCSQDITLSNKTTECKKLTLEKGCIIHARN 320			
QY	297 EMPADPLSLADVESEKOVCKNYAEAKOVFLGMLFYEARRRHPDYSVLLRLAKTYET 356			
DB	321 EKPEGLSPNLNFFLDRODFNPFSSGCKNFILASFVHEYSRRHPOLAVSYILVARQYEL 380			
QY	357 LEKCAADPHHCYAKVDEFRPLVEEPQNLKONCELEFEOLEKCFKFOALVRYTKVP 416			
DB	381 LEKCFQTEPNLEQCKGEEELQYIOESQALAKKSCGLFKQKGEYLYOALVATKPA 440			
QY	417 QVSTPLVEVSHNLKGVSKCKHPBEAKRMPAEDYLSVLNQLCVLHEKTPVSDRYTC 476			
DB	441 QUTSSELMALITPKMATATATCCQSLSEDKILACGEBAAOIIIGHLCIRHEMPVNGVQC 500			
QY	477 CTESLVNRRPCFSALEVDETYVPKFEPAETFTFHADICTLSEKROIKQIALVELYHK 536			
DB	501 CTSSYANRRPCFSSLYVDETYVPAPFSDKFTFHKDLCOAGVALQTMKOEPLMLVYK 560			
QY	537 PKATKEOLKAVMDPAFAVEKCKKADDETCTGAEGSKLVASQALGL 585			
DB	561 POLTEOLEAVYADPSGLLEKCCQGOEDEVCFAEKGLISTKTRALGV 609			

Search completed: July 22, 2003, 11:44:04

Job time : 17 secs

OY 1 DAKSSEVAPFDIGENKALVIAPPAPOYLQOCGEEDHVLNVNEVEFATCYADDSAE 60
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Dd 25 DAKSSIAIRPFKDLGEOHFKGVLVAISFOLOKCYEBEVALNVNEVDFAATCYADDSAE 84
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
OY 61 NCBKSJLHTFGDKLTAVATIRETYGBMADCCKAPEPENEECFLHKDNPMLPYLVAREV 120
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Dd 85 NCOKSJHLTGOKLCAIPLTRDSYSELDDCCAKKRPERNECFLNKHODHPRLPFRVPDA 144
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
OY 121 DVMC7AFHDNETFLKKYLTARHPPIFYAVDELLEFKARKAAPTCCCAADKAACLP 180
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

DB 145 EAMCTSFQENAVTENGHYLHEVARRHPYFAPELLYAEKYSALIMECCGGEADKACITP 204
QY 181 KLDELROGKASAKORLTKCASLOKFGERAFAKAMAVARLSOREPFAEVAESKLVTDLTK 240
DB 205 KLDALKEKALASVNOBLKCSLSLOFGORAFKAMAVAKMSQKFPRADEITKLTDLTK 264
QY 241 VHTCECHDLECCADRADLAKYICENODSISIKLECECEKPLLEKSHCIAEVENDEMPA 300
DB 265 LEECHCHDLECCADRADLAKYICENODSISIKLOACCDKPLVKKSHCSLSEVENDIPA 324
QY 301 DLPSTAAOFVESKOVCKNVAEAKDVFLGMFLYEVARRHPDVSVLLRLAKYETTLKRC 360
DB 325 DLPSTAAOFVEDEKCKNVAEAKDVFLGMFLYEVARRHPDVSVALLLAKYETTLKRC 384
QY 361 CAADPHCEYAKVFEFFPVEEPONLIRKONCELEGEYFONALLVRYTKKPOYST 420
DB 385 CAADPHCEYAKVFEFFPVEEPONLIRKONCELEGEYFONALLVRYTKKPOYST 444
QY 421 PTLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 480
DB 445 PTLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 504
QY 481 LVNRPCFSALEVDITYPKFENAFETPFHADICLSKEROKKOTLVLVYKHKPKAT 540
DB 505 VVERPCFSALPVDITYPKFENAFETPFHADICLSKEROKKOTLVLVYKHKPKAT 564
QY 541 KEOLKAVMDPFAEVEKCKCKADKDETCFAEESKRLVAASQAAL 583
DB 565 GPQLRTVGEFTAFALDKCKKAEKDEKACFSEDSKPLVASSQAAL 607

RESULT 2

ID 08UW05 PRELIMINARY; PRT: 626 AA.
AC 08UW05:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma maculatum (spotted salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF217183; AAL56646.1; -
DR InterPro: IPR00264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum_albumin. 1.
DR SMART: SM00103; ALBUMIN. 3.
DR PROSITE: PS00212; ALBUMIN; UNKNOWN_2.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 626 SERUM ALBUMIN.
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 1e-86; Indels 6; Gaps 3;
Matches 237; Conservative 110; Mismatches 238;

QY 1 DAHKESEVARRFDL-----GEENFKALVLAIAFOYLQOCFEDHVLVNEVTEFACTVAD 56
DB 28 EGHVONPHILIGDILPMIGVNSKGLVLAASQMLPLCPYEHDLORVEDVYQNALDLCKAG 87
QY 57 ESAENCKDSLHTLFSGDKLCTVATLREYVGENADCAKQPERNECFLOHKDND-PNLPRL 115

DB 88 ARHANCASPMYIILDELCKRPENAEKYPFHQEOCKKEDPERHKCFVSHKNAHEBLTKY 147
QY 116 VREPDVWCTAFHONDEFLLKRYLLEIARRHHPYFAPELLYAEKYSALIMECCGGEADKACITP 175
DB 148 VRPAEQICDHAERKGLLARYTFLMILGPHMILPILGFAOFRDGIIVSHCKDQVETA 207
QY 176 A-CLIPKIDELREDECKASSAKORLTKCASLOKFGERAFAKAMAVARLSOREPFAEVAESKL 234
DB 208 GOCFNDKRPENKQVEYEAVALQKNCYILDOFKERALVAAYAVASQKFLPLASFEVNOI 267
QY 235 VTDLTKVHTCECHDLECCADRADLAKYICENODSISIKLECECEKPLLEKSHCIAEVE 294
DB 268 VPDVHNLHQTCGSGMMAKMLERKILTKICEKDEBLTHLKECCDKPLLESACTIITLP 327
QY 295 NDEMPADLPSLAADVESKOVCKNVAEAKDVFLGMFLYEVARRHPDVSVALLLAKYETTLKRC 354
DB 328 NDQKPADLSPKVPYHIDPEVCKLYTEGDTFMGFLYECARRHODYSPENLLRSGSYE 387
QY 355 TLEKCAADPHCEYAKVFEFFPVEEPONLIRKONCELEGEYFONALLVRYTKK 414
DB 388 EFLKCCAAEGHNECLAKTEESLKEIESSVTLTKNGALDKLSYLFQNLIRKYAR 447
QY 415 VPQSTPPLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTK 474
DB 448 MPALSEOSILRTKSMITIGECCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTK 507
QY 475 KCCSTLVNRPCFSALEVDITYPKFENAFETPFHADICLSKEROKKOTLVLVYKHKPKAT 534
DB 508 OCCSHLSQTPCFALPYVDITYPKFENAFETPFHADICLSKEROKKOTLVLVYKHKPKAT 567
QY 535 HKPKATKEOLKAVMDPFAEVEKCKCKADKDETCFAEESKRLVAASQAAL 583
DB 568 QYRHWTDQQLTKCYVNFVPMVDOCKADNHNCFALLEGAKLIDLACKAILAV 618

RESULT 3

ID 08UW06 PRELIMINARY; PRT: 624 AA.
AC 08UW06:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RC Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF217182; AAL56645.1; -
DR InterPro: IPR00264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR ProDom: PD002486; Serum_albumin. 1.
DR SMART: SM00103; ALBUMIN. 3.
DR PROSITE: PS00212; ALBUMIN; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4955EF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 7.1e-75; Indels 12; Gaps 5;
Matches 214; Conservative 102; Mismatches 240;

QY 14 LGSENFKALVLAIAFOYLQOCFEDHVLVNEVTEFACTVADSAENCKDSLHTLFSGDKL 73

```

Db 45 IGVENAKALAMALFSGMLSKCPHHEQVQRVRYNMDIADLCRGAHGDCKGCVMTITLNE 104
QY 74 ICTVATILRETYGEMADOCACOEPRNECCFLQHKDNP-NLPRILVREPVDMCTAFHDNEE 132
Db 105 ICTPENPEKYPHHEGCCCKEDERKCFLEHSTDPKETEYKSPEDICDHENRD 164
QY 133 TELKKLYELIARRHPFYFABELLFFAKRYKAFTTECCQ--AADKACILPKLDELREGKA 191
Db 165 EFLGHYIHKVASSHTMYPRALISFTLHFDGIVSHCCDEATVGQCSKEMPRAHKEVEH 224
QY 192 SSKAKOLKACSLQKFGERRAFKANAVALRSORPKAFAEYSKLVTDLTIKYHTECCGDL 251
Db 225 VCAVQKHNCTIILQNFERRALRASKAANACSKPRHASENVQRLTDGIVLHQTCGCGDDM 284
QY 252 ECADDDADLAKYICENODSISSKLECCCKPILKESHCIAEVENDEMPADLPISLADFE 311
Db 285 ACAAEKMKLTQYCEK-----KKCCKEPVLEKSCVIRLPNDEKRALSPSEVRKYFD 336
QY 312 SKDCKNYAEAKDVFLGMFLYEXARRHPDYSVLLRLAKTYETLLEKCAADPHECYA 371
Db 337 DPEVCKRFKEGDAFNGRFLCDYAKIHPSHAEILNRIASGLEKAYKTCAGAHNCCIA 396
QY 372 KYVDEKPLVEEPQNLKONCELFEOIGEEKFONALLVRYTKVPOVSTPLVEVSRNLG 431
Db 397 KEKETLRHELEAKSKTKLKTGALKEKGPYHFNQIMIVRYTGILPRSSDAFLLYITKTLT 456
QY 432 KVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPTV-SDRYTKCCTESLVNRRPESA 490
Db 457 NIOKCKCLPEDOQMCSEGLGMVFRQIC-QOKKTPFENEKLAHCKDLSFTTPEFAA 515
QY 491 LEVDETVYPAEFNAETFTFHADICTLSEKEROIKQOTALVELVKKPRATEOLKAVMD 550
Db 516 LVYDETVYVPAVAESPFENDEFCTPSEADLOAKKQTFMLHLVTRTHKIIDQYKTISEK 575
QY 551 FAFVKECKCADDKETCFPAEKGKLYAA 578
Db 576 FLAMGCGCCAKADQNECFATEGAKLYEA 603

```

RESULT 4

Q9YGH6 PRELIMINARY; PRT; 603 AA.

AC Q9YGH6; 01-MAY-1999 (TREMblrel. 10, Created)

DT 01-MAY-1999 (TREMblrel. 10, last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, last annotation update)

DE Serum albumin precursor (Fragment).

GN ALB.

OS Rana shqiperica.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.

OX NCBI_TaxID=44326;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BUSHAT; TISSUE-LIVER;

RA Uzzell T., Holtz H.;

RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in frog albumins.";

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U40452; AAD09358.1; -.

DR HSSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot. 3.

DR PRINTS; PRO0802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KM Signal.

FT NON TER 1

FT SIGNAL <1

FT CHAIN 24

FT SEQUENCE 603 AA; 69293 MW; 3403723FA010C99 CRC64;

```

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 8.3e-65;
Matches 190; Conservatve 119; Mismatches 241; Indels 16; Gaps 3;

```

QY 12 KDIGENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTYCVADEAENCKSLHTLFG 71

Db 37 KAVGKPAVEKLVLMVAODEKESCLDENLKVQAKIIEAVDNCCKHPEACCKRALETLYH 96

QY 72 DKICVATILRETYGEMADOCACOEPRNECCFLQHKDNP-LPRILVREPVDMCTAFHDNEE 125

Db 97 DIVCKEDIDQILYPMTECCGKAERTKCFYHRE-----VREEYKIPNIESCK 148

QY 126 AFHDNETELKKLYELIARRHPFYFABELLFFAKRYKAFTTECCQADKACILPKLDEL 185

Db 149 EKHENQRAFSYLSLNIAKHSKLYPRAVLGFIQYNEITTECCADBDKACGEMRPQY 208

QY 186 RDEKASSAKOLKACSLQKFGERRAFKANAVALRSORPKAFAEYSKLVTDLTIKYHTEC 245

Db 209 KKLTVNLEDKHKOCRYLKEFPERSQALTLVQVSORFGNAKYDVEKVTIEIHLNEDC 268

QY 246 CHGDLLECADRDADLAKYICENODSISSKLECCCKPILKESHCIAEVENDEMPADLP 305

Db 269 CKGDVAECMTERMATEHICLAKELSSKLSLDCAGVLERTPTCLALPNEE--PDLP 326

QY 306 AADFEVSKDCKNYAEAKDVFLGMFLYEXARRHPDYSVLLRLAKTYETLLEKCAAD 365

Db 327 LKREYDEHNCENYQDKRYTLHFNHYDRSHQESSPOGLAVRSFEMLEKCCASAN 386

QY 366 PHCYAKVDEKPLVEEPQNLKONCELFEOIGEEKFONALLVRYTKVPOVSTPLVE 425

Db 387 SAECLDAPRLTBAALKENEISKONGALEKIGFNDFYIOLLVRYFGKMPQYATQTLVE 446

QY 426 VSRNLKGVSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPTVSDRYTKCCTESLVNRR 485

Db 447 LTRMAKIGVYCGGLPNKKQPCAEKIDLTLLGEMEREKRTINDVHVCDSYANRR 506

QY 486 PCFSALEVDETVYPAEFNAETFTFHADICTLSEKEROIKQOTALVELVKKPRATEOLK 545

Db 507 PCFTTKLGPYANVAPWDESKLHFTADMCGSADDLQKTLVLYLVEIKKPKCGKEKLT 566

QY 546 AVMDFAFVKECKCADDKETCFPAE 571

Db 567 EVIESFRKTVVECCAENQOACFDEK 592

RESULT 5

Q91134 PRELIMINARY; PRT; 614 AA.

AC Q91134; 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, last annotation update)

DE Cobra serum albumin.

OS Naja naja (Indian cobra).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; OC Elapidae; Elapinae; Naja.

OX NCBI_TaxID=35670;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-LIVER;

RX MEDLINE=96145734; Pubmed=8561913;

RA Wang X., Hansen H., Havsteen B.;

RT "Evidence of the coevolution of snake toxin and its endogenous antitoxin. Cloning, sequence and expression of a serum albumin cDNA of the chinese cobra.";

RT RT Biol. Chem. Hoppe-seyler 376:545-553(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-LIVER;

RX MEDLINE=9343893; Pubmed=8343135;

RA Shao J., Shen H., Havsteen B.;

RT "Purification, characterization and binding interactions of the

RT Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP.
 RL Biochem. J. 293:559-566(1993).
 DR EMBL; X78598; CAA55333.1;
 DR HSSP; P02768; 1E7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 614 AA; 69798 MW; 3DB2D3CC4BD8CBFD CRC64;

Query Match 29.9%; Score 928.5; DB 13; Length 614;
 Best Local Similarity 32.8%; Pred. No. 9e-63;
 Matches 184; Conservative 119; Mismatches 255; Indels 3; Gaps 3;

18 NKAFLVLAFAOYLQCPEDHVKLVNEVEFAKTVADSE-SAEKCDKSLHTLFGDKICT 76
 45 DGAFLTLVLTQVFNATLEDLKLKSAETIELHKCVASFSFSDPPCTKPLGIVFLDVLC 104
 77 VALLRETYGEMADCCAKOEPERNCEFLQHKDNP-NLPRLYRPEVDVMTAFHNEETFL 135
 105 NEFSNKYG-INDCCAKADPDRECEVLSHKSTSTGTSIFPVHNAEACQAFQNDSDVL 163
 136 KKLLEIARRHPYFAPELLEFAKRYKAFTCCQAADKACLLPKLDELDEGRKASSAK 195
 164 AQLIELSRYPALSVVILESTKTYKLTLECCAEADKACIHEKATEAKKFRPEIMEE 223
 196 QRLKASLOKGEGERAFKAAVAARLSORPKAEFAVSKLVTDLTGVHTECCGDLLECAD 255
 224 QETTCYNLKKYKDKLALKFLETHEKRYNAKLETTIGAEVVAHIEETICKGSDVILV 283
 256 DRADLAKYICENODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSLADEYSKDV 315
 284 DRAALSYVCEHKDAISSVGHCCSEKPLEYERPNCLATLANDARSDLPSPSEILLKEIEA 343
 316 CKNVLEADVFLGMFLYEARRHPDYSVVLLRLAKTYETTLKCCAAADPHECTAKYFD 375
 344 CTTYEORENKESFLTLTRNHPKLSKIDLEILYKYEKLEECQSEHNVQCLHGEO 403
 376 EFKPLVEEPONLIKONCELEGEYKFNQALLVRYTKKPOVSPPTLVEYSRNLGKYS 435
 404 VFKLITIKINEVKNCSYKELGDFTFNEFLVYSRMPAPLPSLIELTEKYGKVAE 463
 436 KCKHPEAKRMPCADYLSVNLQCLVLEKTPVSDRYTKCTESLVNRRPFSALAEVDE 495
 464 KCCNDSNHQVSCALENTDKVMSICKYHNKHFINDQICHCCNSFISMECISNLGPDL 523
 496 TYVPEFAETPTFIADICTLSEKROIKKOTALVELYKHKRKAKEOLKAVMDFAFV 555
 524 SFVPTFNPKTMDNEPKLSTSEDTVOKSKGLSELVSKSPNISSEELATILTFREIQ 583
 556 EKCCAKADKETCFEAEGKKLV 576
 584 KLCFAENKKECFDKKGEMV 604

RESULT 6

ID 08R0J9 PRELIMINARY; PRT; 417 AA.
 AC 08R0J9;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 47.7 kDa protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC026681; AAH26681.1;
 KW Hypothetical protein.
 FT NON-TEP
 SQ SEQUENCE 417 AA; 47667 MW; AF84FF0A76A92B3 CRC64;

Query Match 23.0%; Score 713; DB 11; Length 417;
 Best Local Similarity 36.4%; Pred. No. 1.7e-46;
 Matches 143; Conservative 72; Mismatches 170; Indels 8; Gaps 3;

16 EENFALVLAFAOYLQCPEDHVKLVNEVEFAKTVADSESAENCKSLHTLFGDKLC 75
 27 DENTAYLIIAISOYVQASFEDETVLVKVMIDYDRQWADWTLPESKTTADAIQDMIC 86
 76 TVATLRETYGEMADCCAKOEPERNCEFLQHKDNP-NLPRLYRPEVDVMTAFHNEE 131
 87 DMGILPQKH-NFSHCCGAGAPFRRLCFYNNKANGFLPPPTL---DPEECQAYKKN 142
 132 ETEFLKLYEIAARRHPYFAPELLEFAKRYKAFTCCQAADKACLLPKLDELDEGRK 191
 143 ESFLHLYMEVARRNPFEAPVLLAVAAWFEEAATTCCEOQKATCFQAKAPITQYLKA 202
 192 SSARQRLKASLOKGEGERAFKAAVAARLSORPKAEFAVSKLVTDLTGVHTECCGDL 251
 203 SSSYORNVCGALIKFGPVVLSINVAVESKRPKIGFDLTTLLEDSVSMTEGCEGDV 262
 252 ECADRADLAKYICENODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSLADEVE 311
 263 HCIRQSOVNVNICKSDKDISKIVCEKKTLEBEACILANNKDRREGSLRAKFTTE 322
 312 SKDVCKNTAEAKDVLGMFLYEARRHPDYSVVLLRLAKTYETTLKCCAAADPHECTA 371
 323 SENVQERDSDPDCKFEAFIEYYSRRHPDLSPPELLRITIKVMDLEDCSSRENPAQCYR 382
 372 KYDEKPLVEEPONLIKONCELEGEYKFNQALLVRYTKKPOVSPPTLVEYSRNLGKYS 404
 383 HVEDKFNETTORSLAMVQDECKQFOLGDTIQ 415

RESULT 7

ID 09W6F5 PRELIMINARY; PRT; 484 AA.
 AC 09W6F5;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Vitamin-D binding protein.
 GN VTDB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Burnside J., Sofer L.;
 RT "VTDB-chicken vitamin-D binding protein precursor";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF121350; AAD2830.1;
 DR InterPro: IPR001230; Prenyl_site.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 2.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
 SQ SEQUENCE 484 AA; 53686 MW; 2C96758210ECD0F7 CRC64;

Query Match 12.1%; Score 376.5; DB 13; Length 484;
 Best Local Similarity 23.3%; Pred. No. 1e-20;
 Matches 103; Conservative 89; Mismatches 230; Indels 21; Gaps 10;

QY 1 DAHSEVAHREKDLGEENFKALVILFAOYLQOCPEFEDHVLVNEVEFEAKTVADESAE 60
 DB 22 DYEDKONCENLMIKGEDEFSLSLIVSRKSSSPFEVONLVNEVSLTECCAGADP 81
 QY 61 NCDKSLHTLFEDKLCYATLTRETYGEMADCCAKOPEPNEFLOHKDNPMLPRLVREY 120
 DB 82 TCYSTRSELSVKSCSESDAPPVHPGPECTCKEGLKLMALSHOPEFPYVEPTN 141
 QY 121 DVMCTAFHNEETFLKTYELIARHHPYFVAPELLFFAKRYKAAPTECCOADAACLLP 180
 DB 142 DEICFAFRDPKGRADQFLYESSNYGAPRLVLAITKNTLSMGSCCTSANPVYCVK 201
 QY 181 KLDELREGKASAKQRLKASLOKFGERAFAKANAVALSORFPAEPAEYSKLVTDLTK 240
 DB 202 ERLQMKHLSTLTMSNRV-CSQYAAVGKESRLSLIKLAQKVPANLEVLPLAEDPTE 260
 QY 241 VHTECCHDLEEC-ADDRADLAKYICENODSISKLKCCCKPFLKSHCIAEYNDM- 298
 DB 261 ILSRCESTSEDCMASSELPEHTIKQNLKSKNSFECCO-----ENTPMN 307
 QY 299 -----PADLPSELADPVESEKDYC-KNYAEAKDVELGMFLYFARHHPDYSVL 345
 DB 308 IFMCTFPAAPBPIQLP-AIKLPKGLDGCOSTTQAMD---QTFELSRRTQVPEVFL 361
 QY 346 LILRLAKYETTLERKCCAAADPHECYAKVDEFKPLVE-EPONLIKQNCLEFEOLGEYKF- 403
 DB 362 SKVLEPLTK-FLRECCDTQDSVAC---FSTQSPILKRLQSLFIEKGEMCADISENFT 416
 QY 404 --QNALVRYTKKYPQVSTPFLVEYSRNLGKVGSKCC 438
 DB 417 EYKKKLAERLRTKPTNTPSPALCKDMVEKHSDFASKCC 453

RESULT 10

ID 042279 PRELIMINARY; PRT: 551 AA.
 AC 042279;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Serum albumin AS (Fragment).
 GN AS.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX Nchi_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
 RC TISSUE-LARVAL LIVER;
 RX MEDLINE=98428063; PubMed=9755481;
 RA Filosa M.F., Adam I., Robson P., Helnig J.A., Smith K., Keeley F.W.,
 Youson J.H.;
 RT "Partial clone of the gene for AS protein of the lamprey Petromyzon
 marinus, a member of the albumin supergene family whose expression is
 restricted to the larval and metamorphic phases of the life cycle.";
 RL J. Exp. Zool. 282:301-309(1998).
 RN [2]
 RP DEVELOPMENTAL STAGE.
 RX PubMed=10998585;
 RA Dahnis M.H., Filosa M.F., Youson J.H.;
 RT "An albumin-like protein in the serum of non-parasitic brook lamprey
 (Lampetra appendix) is restricted to preadult phases of the life cycle
 in contrast to the parasitic species Petromyzon marinus.";
 RL Comp. Biochem. Physiol. 127B:251-260(2000)
 CC -I- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -I- DEVELOPMENTAL STAGE: LARVAL (AMMOCETE), METAMORPHOSING, AND
 CC JUVENILE INDIVIDUALS, BUT NOT IN SEXUALLY MATURE ADULTS.
 CC -I- DOMAIN COMPOSED OF AT LEAST THREE HOMOLOGOUS DOMAINS.

CC -I- MISCELLANEOUS: IN THE SEA LAMPREY, THERE ARE TWO FORMS OF ALBUMIN,
 AS AND SDS-1.
 CC -I- SIMILARITY: BELONGS TO THE ALB/AFB/VDB FAMILY.
 DR EMBL: AF031134; AAC63407.1; -
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot.3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin.2.
 DR SMART: SM00103; ALBUMIN.2.
 DR PROSITE: PS00212; ALBUMIN.2.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Glycoprotein.
 FT NON_TER 1 1
 FT CHAIN 1 551
 FT DOMAIN <1 540 SERUM ALBUMIN AS.
 FT REPEAT <1 127 3 X APPROXIMATE REPEATS.
 FT REPEAT 146 322 1.
 FT REPEAT 341 540 2.
 FT REPEAT 341 540 3.
 FT DISULFD 5 14 BY SIMILARITY.
 FT DISULFD 27 40 BY SIMILARITY.
 FT DISULFD 39 51 BY SIMILARITY.
 FT DISULFD 73 118 BY SIMILARITY.
 FT DISULFD 117 126 BY SIMILARITY.
 FT DISULFD 149 195 BY SIMILARITY.
 FT DISULFD 194 203 BY SIMILARITY.
 FT DISULFD 216 232 BY SIMILARITY.
 FT DISULFD 231 242 BY SIMILARITY.
 FT DISULFD 266 311 BY SIMILARITY.
 FT DISULFD 310 321 BY SIMILARITY.
 FT DISULFD 344 390 BY SIMILARITY.
 FT DISULFD 389 396 BY SIMILARITY.
 FT DISULFD 409 425 BY SIMILARITY.
 FT DISULFD 424 435 BY SIMILARITY.
 FT DISULFD 485 531 BY SIMILARITY.
 FT DISULFD 530 539 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 551 AA; 61455 MW; 2A92404C824A555 CRC64;

Query Match 11.8%; Score 365; DB 13; Length 551;

Best local similarity 23.0%; Pred. No. 9.3e-20;

Matches 134; Conservative 86; Mismatches 264; Indels 98; Gaps 17;

QY 53 CVADESANCDSKSLTTLFGDKLCYATLTRETYGEMADCCA-KOEPERNECFLOHKDNP 111
 DB 5 CCGKEMAGCILLHRRYLQFDELCEGV--SIPSAACCSLANEBRADCLVSLR-GNLS 60
 QY 112 LPRLVREVDVWCIAFHD---NEETFLKTYELIARHHPYFVAPELLFFAKRYKAATTE 167
 DB 61 IHSVPLAPASQLC--HRRMKSHESF-ASLIMERGRHPRADSOVELEAERFSKIGDA 116
 QY 168 CCOADRAKACLLPKLDELDEGKASAKQRLKASLOKFGERAFAKANAVALSORFPAE 227
 DB 117 CCDLADEKECITTRGREALHOEVSAYADAOLCSSLALGAKFLGRVNVLFSGQRAPAT 176
 QY 228 FAEYSKLTDLTKVHTCC-HGDLLEC-ADDRADLAKYICENODSIS--SKLKECCERPL 283
 DB 177 FDQISKLSHRFHSVAQOTCCGEGWSPGCFABQNHILHDEMCDMELSHVPMAMACQISG 236
 QY 284 LEKSHICAEVENDERPADLPSELADPVESEKDYCKNYAKAVFLGMFLYFARHHPDYSV 343
 DB 237 SARAKCMETIRGRKVLDD--VALARFQGHVCCQANAPQELLGRMLYEFGRRTDSV 293
 QY 344 VLLRLAKYETTLERKCCAA--ADPHECYAKVDEFKPLVEEPONLIKQNCLEFEOLGEY 401
 DB 294 GEAKKITIEMWDYKDCAGNHSQACLVSKKALISVKTIEBQAKSHIKIEOJOKDQHE 353
 QY 402 KFONALVRYTKKYPQVSTPFLVEYSRNLGKVGSKCCNHPAKRMPGCAEDYLVVNLQ 461
 DB 354 VFEKVLDFALIEARTISLDKRVEFARHRYTHAIRCAN---QANCLLDENLHLSISLC 409
 QY 462 VLHKTVPVSDVTKCCTSLVNNRRCFSALEVDYETVYKFNALFTFHADICHTSEKER 521

```

Db 410 SDSLAAHDGYKCC-----RLAESEVSCVHEHERAHEAR 447
QY 522 QIKK-QTALVELVKHKPKATKPOLKAVMDFAA----- 553
Db 448 ATEEVENICKERVENQAKV--EAVEAVEPFPFEBGANSCLFROLPGYLORLLYKAAH 505
QY 554 -----FVE---KCKADKETCEAECKKL 575
Db 506 QAPAGVDHSRIRLOVHHFEVYAKCCORAVDKECSHETKEM 547

RESULT 11
Q90WZ8 PRELIMINARY: PRT: 122 AA.
AC Q90WZ8:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serum albumin (Fragment).
OS Larus argentatus (Herring gull).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Charadriiformes; Laridae; Larus.
OX NCBI_TaxID=35669;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzen A., Casley W.L., Moon T.W.;
RT "Development of an RT-PCR Bioassay for Avian Vitellogenin mRNA.";
RL Toxicol. Appl. Pharmacol. 0:0-0(2001).
DR EMBL: AY045725; AA01533.1;
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13805 MW; 93C644A0B120EF93 CRC64;

Query Match
Best Local Similarity 47.1%; Score 331; DB 13; Length 122;
Matches 57; Conservative 23; Mismatches 41; Indels 0; Gaps 0;

QY 453 YLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTFPA 511
Db 1 YLSIVIQDCRQETTPVNDVSHCCSDSYAVRRPCTFMGVDTKYPPAPFPEMFSPDE 60

QY 512 DICTLSEKEROIKQTALVELVKHKPKATKPOLKAVMDFAAFAVEKCKADKCTCFAE 571
Db 61 KICTAPPAEDELGQKLLINLKKRPQMTIEQIKTADFTAMVDKCKQSDIETCFGE 120

QY 572 G 572
Db 121 G 121

RESULT 12
Q90WZ6 PRELIMINARY: PRT: 123 AA.
AC Q90WZ6:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serum albumin (Fragment).
OS Poephila guttata (Zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Passeriformes; Estrildidae;
OC Estrildinae; Taeniopygia.
OX NCBI_TaxID=59729;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzen A., Casley W.L., Moon T.W.;
RT "Development of an RT-PCR Bioassay for Avian Vitellogenin mRNA.";
RL Toxicol. Appl. Pharmacol. 0:0-0(2001).
DR EMBL: AY045727; AA01535.1;
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.

```

```

FT NON_TER 1
FT NON_TER 123
SQ SEQUENCE 123 AA; 13824 MW; B18F83BC2E194F29 CRC64;

Query Match
Best Local Similarity 43.0%; Score 290; DB 13; Length 123;
Matches 52; Conservative 22; Mismatches 47; Indels 0; Gaps 0;

QY 453 YLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTFPA 511
Db 1 YLTVLIEDCKKQESTPVNDVSHCCNDYSPKRCFTFMGDTKYVPPAPFTLFDDE 60

QY 512 DICTLSEKEROIKQTALVELVKHKPKATKPOLKAVMDFAAFAVEKCKADKCTCFAE 571
Db 61 KICKAPPAERAGELKLVNLYKRPQMTIEQIKTBEFTAMMECKCKPVEGCLGE 120

QY 572 G 572
Db 121 G 121

RESULT 13
Q63205 PRELIMINARY: PRT: 135 AA.
AC Q63205:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Messenger RNA for rat alpha-fetoprotein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-81006964; Pubmed-6157690;
RA Innis M.A., Miller D.L.;
RT "Alpha-fetoprotein gene expression. Partial DNA sequence and COOH-terminal homology to albumin.";
RL J. Biol. Chem. 255:8994-8996(1980).
DR EMBL: V01236; CAA24546.1;
DR HSSP: P02768; IE7B.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.
DR SMART: SM00103; ALBUMIN; 1.
FT NON_TER 1
FT NON_TER 135
SQ SEQUENCE 135 AA; 15040 MW; 0F922DAFAF71EAC4 CRC64;

Query Match
Best Local Similarity 41.0%; Score 264; DB 11; Length 135;
Matches 50; Conservative 20; Mismatches 52; Indels 0; Gaps 0;

QY 450 EDYLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTF 509
Db 1 EGLADIYIGLCLRHANVNSGINHCSSSNRLCTISFLRETTYVPLPFSATNSS 60

QY 510 HADICLSEKEROIKQTALVELVKHKPKATKPOLKAVMDFAAFAVEKCKADKCTCF 569
Db 61 TRVLCOAGRAPQMTKQELLINLVKQKPMTEBOAAVYADPSGLIECKCKDDQDQACFA 120

QY 570 EE 571
Db 121 KE 122

RESULT 14
Q9JMX8 PRELIMINARY: PRT: 1723 AA.
AC Q9JMX8:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Cag-Y.

```

GN CAG-Y.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxId=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CCUG 17874;
RA MEDLINE=97121442; PubMed=8962108;
RA Gensini S., Lange C., Xiang Z., Crabtree J.E., Ghisara P.,
RA Borodovsky M., Rappaport R., Covacci A.;
RA "cag, a pathogenicity island of Helicobacter pylori, encodes type I-
RT specific and disease-associated virulence factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14648-14653(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CCUG 17874;
RA MEDLINE=20150112; PubMed=10684850;
RA Covacci A., Rappaport R.;
RA "Tyrosine-phosphorylated bacterial proteins: Trojan horses for the
RT host cell.";
RL J. Exp. Med. 191:587-592(2000).
DR EMBL: AF282852; AAF80198.1; -.
SQ SEQUENCE 1723 AA; 196048 MW; 0FAC456B76622801 CRC64;

Query Match 6.1%; Score 188; DB 2; Length 1723;
Best Local Similarity 22.9%; Pred. No. 1.2e-05;
Matches 157; Conservative 88; Mismatches 254; Indels 188; Gaps 37;

OY 16 EENFALVLIFAOYLQCPEDHYKLVNEYTEPAKTCVADESAENC-----DKSLHTLF 70
DB 351 EERIKCLDLIDENLKLKSLNQKQVALDCLKNKKT--DEERNECKLINDPEIRKRF 407
OY 71 GDKLTVAITLRETYGEMADCC--AKQEPERNECFLOHDDNPNRLRYPREVDVWCTAFH 128
DB 408 RKEL-----GLOKELOEYDKICLNKTAETAEKNECKLGLSKE--AIERLQOALDCLKNKKT 461
OY 129 DNEETFLKLYLEIARRHPYFAPELL--FFAKRYKAFTCCQADRAAC---LLPK-- 181
DB 462 DEERN-----ECKLNIPDOLQKELLADMSYKAKDCVSKARNEKEKECEKLLTPPEAR 514
OY 182 -----LDELDEKASAKORLKAS-----LQKGEAFKAMAVARLSQFPP 224
DB 515 KLEQOVLDCIKN--AKTDERKKCLDLPKQDLSILAKESLAKYKDC---VSQAKT 567
OY 225 KAEEFAVSKLYT-----DLTKVHTEC-----CHGDILLECADRDLD 260
DB 568 EAERKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEAKKECEKLLTPPAKLLLE 627
OY 261 AK-----YI-CENODSISKLEKCE-----KPLLEKS--HCI--AEVEND-----EM 298
DB 628 AKKSVKAYLDVSOAKTEDEKKECEKLLTPPAKLLLEQALDCLKNKTAEDKKRCVDL 687
OY 299 PADLPS--LAAD-----FVESKDVCNVAEADVFLGMFLYEARRHPDYSVLLRLAKYT 352
DB 668 PKDLQKVLAEESVRYVLDVSKARNEAEERKE--CEKLLTPPAK-------LLEBAKK 737
OY 353 YETTLERCCCAA--DPHECTYAKVFDEKFLVEEPOMLIKONCELFQDL--GEYKQON 405
DB 738 SVKAKDCVSRARNKEKEKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEDEKKECE 797
OY 406 ALLVRYTKRVPQVSTPLVE---VSRNLGVGSKCKH--PEAKRMPCAEDYLSVVLN 458
DB 798 KLLTPPAKLLLEBAKESVAYLDVSOAKTEDEKKECEKLLTPPAKLL-----LE 847
OY 459 QL---CVLHEKTPVSDRYTKCTESTLVNRRPCFSALVEDETVYVPEFNAEFTT-----FH 510
DB 848 QOALDCLKNKTAED--KKRCVKDL-----PKDLQKVLAKKSVAY 887
OY 511 ADICTLSEKEROIK--OTALVELVYHKPKATREOLKAVMDPAFVVERCCADKDETC-- 567
DB 888 LDCVSRARNEKEKECEKLLTPPAKLLLEBAKESLKAIKD-----CLSOARNEEERRACEK 943

OY 568 -----FAEBGKVLAA-----SQA 581
DB 944 LITPEARKLLEBAKESVAYLDVSOA 970

RESULT 15
ID Q9ZLV0 PRELIMINARY; PRT; 1819 AA.
AC Q9ZLV0;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CAG island protein.
GN JHP0476.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxId=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Orla-Nielsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RA "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL: AE001481; AAD06047.1; -.
SQ SEQUENCE 1819 AA; 207505 MW; E76762B5A7FA371D CRC64;

Query Match 5.9%; Score 184; DB 16; Length 1819;
Best Local Similarity 21.8%; Pred. No. 2.7e-05;
Matches 150; Conservative 101; Mismatches 278; Indels 158; Gaps 34;

OY 16 EENFALVLIFAOYLQCPEDHYKLVNEYTEPAKTCVADESAENC-----DKSLHTLF 70
DB 371 EERIKCLDLIDENLKLKSLNQKQVALDCLKNKKT--DEERNECKLINDPEIRKRF 427
OY 71 GDKLTVAITLRETYGEMADCC--AKQEPERNECFLOHDDNPNRLRYPREVDVWCTAFH 128
DB 428 RKEL-----ELOKELOEYDKICLNKTAETAEKNECKLGLSKE--AIERLQOALDCLKNKKT 481
OY 129 DNEETFLKLYLEIARRHPYFAPELL--FFAKRYKAFTCCQADRAAC---LLPKLD 183
DB 482 DEERN-----ECKLNIPDOLQKELLADMSYKAKDCVSKARNEKEKECEKLLTPPAK 534
OY 184 ELRDE-----GKASSAKORLKAS-----LQKGEAFKAMAVARLSQFPPAE 227
DB 535 KLEQOVLDCIKNKTDEKKECEKLLTPKQDLSILAKESLAKYKDC---VSQAKTAE 590
OY 228 FAEVSCLYT-----DLTKVHTEC-----CHGDILLECADRDLD 262
DB 591 KKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEAKKECEKLLTPPAKLLLEBAK 650
OY 263 ---YI-CENODSISKLEKCE-----KPLLEKS--HCI--AEVEND-----MPAD 301
DB 651 SVRAYLDVSKARNEAEKECEKLLTPPAKLLLEQALDCLKNKTAEDERRECKLDPKD 710
OY 302 LPS--LAAD-----FVESKDVCNVAEADVFLGMFLYEARRHPDYSVLLRLAKTYET 355
DB 711 LQKVLAEESVRYVLDVSKAKNEAEERKE--CEKLLTPPAK-------LLEBAKKSVK 760
OY 356 TLEKCCAAA--DPHECTYAKVFDEKFLVEEPOMLIKONCELFQDL--GEYKQONLL 408
DB 761 AYKDCVLRARNEKEKECEKLLTPPAKLLLEBSKSVAYLDVSKARNEAEERKECEKLL 820
OY 409 VRYTKRVPQVSTPLVE---VSRNLGVGSKCKH--PEAKRMPCAEDYLSVVLNQ 460
DB 821 TPPAKLLLEBAKESVAYLDVSKARNE--EKEKQCEKLLTPPAKLL--LEESKSVAYLD 878

THIS PAGE BLANK (uspto)